

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
29 August 2002 (29.08.2002)

PCT

(10) International Publication Number  
**WO 02/066627 A1**

(51) International Patent Classification<sup>7</sup>: C12N 9/48,  
15/12, 5/10, C12Q 1/68, G01N 33/50, C12N 15/62

(21) International Application Number: PCT/EP02/01538

(22) International Filing Date: 14 February 2002 (14.02.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/268,863 16 February 2001 (16.02.2001) US

(71) Applicant (for all designated States except US): BAYER  
AKTIENGESELLSCHAFT [DE/DE]; 51368 Lev-  
erkusen (DE).

(72) Inventor; and

(75) Inventor/Applicant (for US only): LIOU, Jiing, Ren  
[CN/US]; 10 Winslow Road, Belmont, MA 02478 (US).

(74) Common Representative: BAYER AKTIENGE-  
SELLSCHAFT; 51368 Leverkusen (DE).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,  
SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,  
VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,  
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent  
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,  
NE, SN, TD, TG).

**Declarations under Rule 4.17:**

- as to applicant's entitlement to apply for and be granted  
a patent (Rule 4.17(ii)) for the following designations AE,  
AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA,  
CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES,  
FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,  
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG,  
MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU,  
SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG,  
UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS,  
MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent  
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent  
(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,  
MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI,  
CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)
- as to the applicant's entitlement to claim the priority of the  
earlier application (Rule 4.17(iii)) for all designations

**Published:**

- with international search report
- before the expiration of the time limit for amending the  
claims and to be republished in the event of receipt of  
amendments

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

(54) Title: REGULATION OF HUMAN DIPEPTIDYL PEPTIDASE 8

(57) Abstract: Reagents that regulate human dipeptidyl peptidase 8 and reagents which bind to human dipeptidyl peptidase 8 gene products can play a role in preventing, ameliorating, or correcting dysfunctions or diseases including, but not limited to, cancer, CNS disorders, and COPD.



WO 02/066627 A1

## **REGULATION OF HUMAN DIPEPTIDYL PEPTIDASE 8**

### **TECHNICAL FIELD OF THE INVENTION**

5

The invention relates to the regulation of human dipeptidyl peptidase 8.

### **BACKGROUND OF THE INVENTION**

10 Dipeptidyl peptidase is responsible for the removal of terminal dipeptides sequentially from polypeptides. There is a need in the art to identify related enzymes, which can be regulated to provide therapeutic effects.

### **SUMMARY OF THE INVENTION**

15

It is an object of the invention to provide reagents and methods of regulating a human dipeptidyl peptidase 8. This and other objects of the invention are provided by one or more of the embodiments described below.

20 One embodiment of the invention is a dipeptidyl peptidase 8 polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 62% identical to  
the amino acid sequence shown in SEQ ID NO: 2 and;

25 the amino acid sequence shown in SEQ ID NO: 2;

Yet another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a dipeptidyl peptidase 8 polypeptide comprising an amino acid sequence selected from  
30 the group consisting of:

amino acid sequences which are at least about 62% identical to  
the amino acid sequence shown in SEQ ID NO: 2 and;  
the amino acid sequence shown in SEQ ID NO: 2;

5 Binding between the test compound and the dipeptidyl peptidase 8 polypeptide is detected. A test compound which binds to the dipeptidyl peptidase 8 polypeptide is thereby identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the activity of the dipeptidyl peptidase 8.

10 Another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a polynucleotide encoding a dipeptidyl peptidase 8 polypeptide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

15 nucleotide sequences which are at least about 50% identical to  
the nucleotide sequence shown in SEQ ID NO: 1 and;  
the nucleotide sequence shown in SEQ ID NO: 1;

20 Binding of the test compound to the polynucleotide is detected. A test compound which binds to the polynucleotide is identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the amount of the dipeptidyl peptidase 8 through interacting with the dipeptidyl peptidase 8 mRNA.

25 Another embodiment of the invention is a method of screening for agents which regulate extracellular matrix degradation. A test compound is contacted with a dipeptidyl peptidase 8 polypeptide comprising an amino acid sequence selected from the group consisting of:

30 amino acid sequences which are at least about 62% identical to  
the amino acid sequence shown in SEQ ID NO: 2 and;  
the amino acid sequence shown in SEQ ID NO: 2;

A dipeptidyl peptidase 8 activity of the polypeptide is detected. A test compound which increases dipeptidyl peptidase 8 activity of the polypeptide relative to dipeptidyl peptidase 8 activity in the absence of the test compound is thereby identified  
5 as a potential agent for increasing extracellular matrix degradation. A test compound which decreases dipeptidyl peptidase 8 activity of the polypeptide relative to dipeptidyl peptidase 8 activity in the absence of the test compound is thereby identified as a potential agent for decreasing extracellular matrix degradation.

10 Even another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a dipeptidyl peptidase 8 product of a polynucleotide which comprises a nucleotide sequence selected from the group consisting of:

15 nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1 and;  
the nucleotide sequence shown in SEQ ID NO: 1;

Binding of the test compound to the dipeptidyl peptidase 8 product is detected. A test  
20 compound which binds to the dipeptidyl peptidase 8 product is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Still another embodiment of the invention is a method of reducing extracellular matrix degradation. A cell is contacted with a reagent which specifically binds to a  
25 polynucleotide encoding a dipeptidyl peptidase 8 polypeptide or the product encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to  
30 the nucleotide sequence shown in SEQ ID NO: 1 and;  
the nucleotide sequence shown in SEQ ID NO: 1;



Dipeptidyl peptidase 8 activity in the cell is thereby decreased.

5 The invention thus provides a human dipeptidyl peptidase 8 that can be used to identify test compounds that may act, for example, as activators or inhibitors at the enzyme's active site. Human dipeptidyl peptidase 8 and fragments thereof also are useful in raising specific antibodies that can block the enzyme and effectively reduce its activity.

10 **BRIEF DESCRIPTION OF THE DRAWINGS**

Fig. 1 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 1).

15 Fig. 2 shows the amino acid sequence deduced from the DNA-sequence of Fig.1 (SEQ ID NO: 2).

Fig. 3 shows the amino acid sequence of a protein identified by trembl|AF221634|AF221634\_1 (SEQ ID NO: 3).

20

Fig. 4 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 4).

25 Fig. 5 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 5).

Fig. 6 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 6).

30 Fig. 7 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 7).

- Fig. 8 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 8).
- 5 Fig. 9 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 9).
- Fig. 10 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 10).
- 10 Fig. 11 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 11).
- Fig. 12 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 12).
- 15 Fig. 13 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 13).
- Fig. 14 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 14).
- 20 Fig. 15 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 15).
- Fig. 16 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 16).
- 25 Fig. 17 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 17).
- 30

Fig. 18 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 18).

5 Fig. 19 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 19).

Fig. 20 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 20).

10 Fig. 21 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 21).

Fig. 22 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 22).

15 Fig. 23 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 23).

20 Fig. 24 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 24).

Fig. 25 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 25).

25 Fig. 26 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 26).

Fig. 27 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 27).

30

Fig. 28 shows the DNA-sequence encoding a dipeptidyl peptidase 8 Polypeptide (SEQ ID NO: 28).

5 Fig. 29 shows the BLASTP - alignment of 181\_Protein (SEQ ID NO: 2) against trembl|AF221634|AF221634\_1 (SEQ ID NO: 3).

Fig. 30 shows the HMMPFAM - alignment of 181\_Protein (SEQ ID NO: 2) against pfam|hmm|DPPIV\_N\_term.

10 Fig. 31 shows the HMMPFAM - alignment of 181\_Protein (SEQ ID NO: 2) against pfam|hmm|Peptidase\_S9.

#### **DETAILED DESCRIPTION OF THE INVENTION**

15 The invention relates to an isolated polynucleotide from the group consisting of:

- a) a polynucleotide encoding a dipeptidyl peptidase 8 polypeptide comprising an amino acid sequence selected from the group consisting of:
  - 20 amino acid sequences which are at least about 62% identical to the amino acid sequence shown in SEQ ID NO: 2 and;
  - the amino acid sequence shown in SEQ ID NO: 2;
- b) a polynucleotide comprising the sequence of SEQ ID NO: 1;
- 25 c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b) and encodes a dipeptidyl peptidase 8 polypeptide;

- d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code and encodes a dipeptidyl peptidase 8 polypeptide; and
- 5 e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d) and encodes a dipeptidyl peptidase 8 polypeptide.

10 Furthermore, it has been discovered by the present applicant that a novel dipeptidyl peptidase 8, particularly a human dipeptidyl peptidase 8, can be used in therapeutic methods to treat cancer, a CNS disorder or COPD.

Human dipeptidyl peptidase 8 comprises the amino acid sequence shown in SEQ ID NO: 2. A coding sequence for human dipeptidyl peptidase 8 is shown in SEQ ID  
15 NO: 1. This sequence is located on chromosome 19. Related ESTs (SEQ ID NOS: 4-28) are expressed in colon; muscle (rhabdomyosarcoma); placenta (chorio-carcinoma); nervous\_tumor; skin (melanotic melanoma); bone marrow (acute myelogenous leukemia); lung (small cell carcinoma); uterus\_tumor; breast; marrow; lung\_normal; adrenal\_gland (adrenal adenoma); and head\_neck.

20 Human dipeptidyl peptidase 8 is 61% identical over 840 amino acids to trembl|AF221634|AF221634\_1 (SEQ ID NO: 3) (FIG. 1). Both pfam and blocks searches confirm this protein's peptidase function based on its homology to dipeptidyl peptidase and prolyl oligopeptidase. In addition, the critical active site  
25 serine, aspartate, and histidine residues are found in the molecule.

Human dipeptidyl peptidase 8 of the invention is expected to be useful for the same purposes as previously identified dipeptidyl peptidase 8 enzymes. Human dipeptidyl peptidase 8 is believed to be useful in therapeutic methods to treat disorders such as  
30 cancer, CNS disorders, and COPD. Human dipeptidyl peptidase 8 also can be used to screen for human dipeptidyl peptidase 8 activators and inhibitors.

### Polypeptides

Human dipeptidyl peptidase 8 polypeptides according to the invention comprise at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, or 892 contiguous amino acids selected from the amino acid sequence shown in SEQ ID NO: 2 or a biologically active variant thereof, as defined below. A dipeptidyl peptidase 8 polypeptide of the invention therefore can be a portion of a dipeptidyl peptidase 8 protein, a full-length dipeptidyl peptidase 8 protein, or a fusion protein comprising all or a portion of a dipeptidyl peptidase 8 protein.

### Biologically Active Variants

Human dipeptidyl peptidase 8 polypeptide variants that are biologically active, e.g., retain a dipeptidyl peptidase activity, also are dipeptidyl peptidase 8 polypeptides. Preferably, naturally or non-naturally occurring dipeptidyl peptidase 8 polypeptide variants have amino acid sequences which are at least about 62, 65, or 70, preferably about 75, 80, 85, 90, 96, 96, 98, or 99% identical to the amino acid sequence shown in SEQ ID NO: 2 or a fragment thereof. Percent identity between a putative dipeptidyl peptidase 8 polypeptide variant and an amino acid sequence of SEQ ID NO: 2 is determined using the Blast2 alignment program (Blosum62, Expect 10, standard genetic codes).

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity of a dipeptidyl peptidase 8 polypeptide can be found using computer programs well known in the art, such as DNASTAR software. Whether an amino acid change results in a biologically active dipeptidyl peptidase 8 polypeptide can readily be determined by assaying for dipeptidyl peptidase activity, as described for example, in Maes *et al.*, Neuropsychopharmacology. 2001 Feb;24(2):130-40; Sentandreu & Toldra, J Agric Food Chem. 2000 Oct;48(10):5014-22; Li *et al.*, Biochem Biophys Res Commun. 2000 Sep 24;276(2):553-8; or Durinx *et al.*, Eur J Biochem. 2000 Sep;267(17):5608-13.

#### Fusion Proteins

Fusion proteins are useful for generating antibodies against dipeptidyl peptidase 8 polypeptide amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins that interact with portions of a dipeptidyl peptidase 8 polypeptide. Protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A dipeptidyl peptidase 8 polypeptide fusion protein comprises two polypeptide segments fused together by means of a peptide bond. The first polypeptide segment comprises at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, or 892 contiguous amino acids of SEQ ID NO: 2 or of a biologically active variant, such as those described above. The first polypeptide segment also can comprise full-length dipeptidyl peptidase 8 protein.

The second polypeptide segment can be a full-length protein or a protein fragment. Proteins commonly used in fusion protein construction include  $\beta$ -galactosidase,  $\beta$ -glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between the dipeptidyl peptidase 8 polypeptide-encoding sequence and the heterologous protein sequence, so that the dipeptidyl peptidase 8 polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two polypeptide segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises coding sequences selected from SEQ ID NO: 1 in proper reading frame with nucleotides encoding the second polypeptide segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

#### Identification of Species Homologs

Species homologs of human dipeptidyl peptidase 8 polypeptide can be obtained using dipeptidyl peptidase 8 polypeptide polynucleotides (described below) to make



suitable probes or primers for screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs which encode homologs of dipeptidyl peptidase 8 polypeptide, and expressing the cDNAs as is known in the art.

5

### Polynucleotides

A dipeptidyl peptidase 8 polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for a dipeptidyl peptidase 8 polypeptide. A coding sequence for human dipeptidyl peptidase 8 is shown in SEQ ID NO: 1.

Degenerate nucleotide sequences encoding human dipeptidyl peptidase 8 polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, 98, or 99% identical to the nucleotide sequence shown in SEQ ID NO: 1 or its complement also are dipeptidyl peptidase 8 polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologs, and variants of dipeptidyl peptidase 8 polynucleotides that encode biologically active dipeptidyl peptidase 8 polypeptides also are dipeptidyl peptidase 8 polynucleotides. Polynucleotide fragments comprising at least 8, 9, 10, 11, 12, 15, 20, or 25 contiguous nucleotides of SEQ ID NO: 1 or its complement also are dipeptidyl peptidase 8 polynucleotides. These fragments can be used, for example, as hybridization probes or as antisense oligonucleotides.

### Identification of Polynucleotide Variants and Homologs

Variants and homologs of the dipeptidyl peptidase 8 polynucleotides described above also are dipeptidyl peptidase 8 polynucleotides. Typically, homologous dipeptidyl

peptidase 8 polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known dipeptidyl peptidase 8 polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions--2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1% SDS, 50°C once, 30 minutes; then 2X SSC, room temperature twice, 10 minutes each--homologous sequences can be identified which contain at most about 25-30% basepair mismatches. More preferably, homologous nucleic acid strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologs of the dipeptidyl peptidase 8 polynucleotides disclosed herein also can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of dipeptidyl peptidase 8 polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the  $T_m$  of a double-stranded DNA decreases by 1-1.5°C with every 1% decrease in homology (Bonner *et al.*, *J. Mol. Biol.* 81, 123 (1973)). Variants of human dipeptidyl peptidase 8 polynucleotides or dipeptidyl peptidase 8 polynucleotides of other species can therefore be identified by hybridizing a putative homologous dipeptidyl peptidase 8 polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO: 1 or the complement thereof to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to dipeptidyl peptidase 8 polynucleotides or their complements following stringent hybridization and/or wash conditions also are dipeptidyl peptidase 8 polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989, at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20°C below the calculated  $T_m$  of the hybrid under study. The  $T_m$  of a hybrid between a dipeptidyl peptidase 8 polynucleotide having a nucleotide sequence shown in SEQ ID NO: 1 or the complement thereof and a polynucleotide sequence which is at least about 50, preferably about 75, 90, 96, or 98% identical to one of those nucleotide sequences can be calculated, for example, using the equation of Bolton and McCarthy, *Proc. Natl. Acad. Sci. U.S.A.* 48, 1390 (1962):

$$T_m = 81.5^\circ\text{C} - 16.6(\log_{10}[\text{Na}^+]) + 0.41(\%G + C) - 0.63(\%\text{formamide}) - 600/l,$$

where  $l$  = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4X SSC at 65°C, or 50% formamide, 4X SSC at 42°C, or 0.5X SSC, 0.1% SDS at 65°C. Highly stringent wash conditions include, for example, 0.2X SSC at 65°C.

#### Preparation of Polynucleotides

A dipeptidyl peptidase 8 polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, or synthesized using an amplification technique, such as the polymerase chain reaction (PCR), or by using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated dipeptidyl peptidase 8 polynucleotides. For example, restriction enzymes and probes can be used to isolate polynucleotide fragments, which comprise dipeptidyl peptidase 8 nucleotide sequences. Isolated polynucleotides are in preparations that are free or at least 70, 80, or 90% free of other molecules.

Human dipeptidyl peptidase 8 cDNA molecules can be made with standard molecular biology techniques, using dipeptidyl peptidase 8 mRNA as a template. Human dipeptidyl peptidase 8 cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as  
5 Sambrook *et al.* (1989). An amplification technique, such as PCR, can be used to obtain additional copies of polynucleotides of the invention, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesize dipeptidyl  
10 peptidase 8 polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a dipeptidyl peptidase 8 polypeptide having, for example, an amino acid sequence shown in SEQ ID NO: 2 or a biologically active variant thereof.

15 Extending Polynucleotides

Various PCR-based methods can be used to extend the nucleic acid sequences disclosed herein to detect upstream sequences such as promoters and regulatory elements. For example, restriction-site PCR uses universal primers to retrieve  
20 unknown sequence adjacent to a known locus (Sarkar, *PCR Methods Applic.* 2, 318-322, 1993). Genomic DNA is first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are  
25 transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (Triglia *et al.*, *Nucleic Acids Res.* 16, 8186, 1988). Primers  
30 can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 22-30

nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72°C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

5

Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom *et al.*, *PCR Methods Applic. 1*, 111-119, 1991). In this method, multiple restriction enzyme digestions and ligations also can be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

10

Another method which can be used to retrieve unknown sequences is that of Parker *et al.*, *Nucleic Acids Res. 19*, 3055-3060, 1991). Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to walk genomic DNA (CLONTECH, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

15

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Randomly-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' non-transcribed regulatory regions.

20

25

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) that are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate

30

software (e.g. GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the entire process from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA that might be present in limited amounts in a particular sample.

#### Obtaining Polypeptides

Human dipeptidyl peptidase 8 polypeptides can be obtained, for example, by purification from human cells, by expression of dipeptidyl peptidase 8 polynucleotides, or by direct chemical synthesis.

#### Protein Purification

Human dipeptidyl peptidase 8 polypeptides can be purified from any cell that expresses the polypeptide, including host cells that have been transfected with dipeptidyl peptidase 8 expression constructs. A purified dipeptidyl peptidase 8 polypeptide is separated from other compounds that normally associate with the dipeptidyl peptidase 8 polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified dipeptidyl peptidase 8 polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis.

#### Expression of Polynucleotides

To express a dipeptidyl peptidase 8 polynucleotide, the polynucleotide can be inserted into an expression vector that contains the necessary elements for the

transcription and translation of the inserted coding sequence. Methods that are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding dipeptidyl peptidase 8 polypeptides and appropriate transcriptional and translational control elements. These methods include *in vitro* re-

5 recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook *et al.* (1989) and in Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.

10 A variety of expression vector/host systems can be utilized to contain and express sequences encoding a dipeptidyl peptidase 8 polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems infected with virus expression vectors

15 (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those non-translated regions of the

20 vector -- enhancers, promoters, 5' and 3' untranslated regions -- which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used. For example, when cloning in

25 bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (*e.g.*, heat shock, RUBISCO, and storage protein genes) or from plant viruses (*e.g.*,

30 viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are

preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a dipeptidyl peptidase 8 polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

5     Bacterial and Yeast Expression Systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the dipeptidyl peptidase 8 polypeptide. For example, when a large quantity of a dipeptidyl peptidase 8 polypeptide is needed for the induction of  
10     antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multi-functional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene). In a BLUESCRIPT vector, a sequence encoding the dipeptidyl peptidase 8 polypeptide can be ligated into the vector in frame with sequences for the amino-terminal  
15     Met and the subsequent 7 residues of  $\beta$ -galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem.* 264, 5503-5509, 1989) or pGEX vectors (Promega, Madison, Wis.) also can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells  
20     by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

25     In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel *et al.* (1989) and Grant *et al.*, *Methods Enzymol.* 153, 516-544, 1987.



Plant and Insect Expression Systems

If plant expression vectors are used, the expression of sequences encoding dipeptidyl  
peptidase 8 polypeptides can be driven by any of a number of promoters. For  
5 example, viral promoters such as the 35S and 19S promoters of CaMV can be used  
alone or in combination with the omega leader sequence from TMV (Takamatsu,  
*EMBO J.* 6, 307-311, 1987). Alternatively, plant promoters such as the small subunit  
of RUBISCO or heat shock promoters can be used (Coruzzi *et al.*, *EMBO J.* 3,  
1671-1680, 1984; Broglie *et al.*, *Science* 224, 838-843, 1984; Winter *et al.*, *Results*  
10 *Probl. Cell Differ.* 17, 85-105, 1991). These constructs can be introduced into plant  
cells by direct DNA transformation or by pathogen-mediated transfection. Such  
techniques are described in a number of generally available reviews (e.g., Hobbs or  
Murray, in MCGRAW HILL YEARBOOK OF SCIENCE AND TECHNOLOGY, McGraw Hill,  
New York, N.Y., pp. 191-196, 1992).

15

An insect system also can be used to express a dipeptidyl peptidase 8 polypeptide.  
For example, in one such system *Autographa californica* nuclear polyhedrosis virus  
(AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells  
or in *Trichoplusia* larvae. Sequences encoding dipeptidyl peptidase 8 polypeptides  
20 can be cloned into a non-essential region of the virus, such as the polyhedrin gene,  
and placed under control of the polyhedrin promoter. Successful insertion of  
dipeptidyl peptidase 8 polypeptides will render the polyhedrin gene inactive and  
produce recombinant virus lacking coat protein. The recombinant viruses can then  
be used to infect *S. frugiperda* cells or *Trichoplusia* larvae in which dipeptidyl  
25 peptidase 8 polypeptides can be expressed (Engelhard *et al.*, *Proc. Nat. Acad. Sci.* 91,  
3224-3227, 1994).

Mammalian Expression Systems

30 A number of viral-based expression systems can be used to express dipeptidyl  
peptidase 8 polypeptides in mammalian host cells. For example, if an adenovirus is

- used as an expression vector, sequences encoding dipeptidyl peptidase 8 polypeptides can be ligated into an adenovirus transcription/translation complex comprising the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome can be used to obtain a viable virus that is capable of
- 5 expressing a dipeptidyl peptidase 8 polypeptide in infected host cells (Logan & Shenk, *Proc. Natl. Acad. Sci.* 81, 3655-3659, 1984). If desired, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.
- 10 Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (*e.g.*, liposomes, polycationic amino polymers, or vesicles).
- 15 Specific initiation signals also can be used to achieve more efficient translation of sequences encoding dipeptidyl peptidase 8 polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a dipeptidyl peptidase 8 polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or
- 20 translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins,
- 25 both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used (see Scharf *et al.*, *Results Probl. Cell Differ.* 20, 125-162, 1994).

Host Cells

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed dipeptidyl peptidase 8 polypeptide in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. Different host cells that have specific cellular machinery and characteristic mechanisms for post-translational activities (*e.g.*, CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express dipeptidyl peptidase 8 polypeptides can be transformed using expression vectors which can contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 1-2 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced dipeptidyl peptidase 8 sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. See, for example, ANIMAL CELL CULTURE, R.I. Freshney, ed., 1986.

Any number of selection systems can be used to recover transformed cell lines.

These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, *Cell* 11, 223-32, 1977) and adenine phosphoribosyltransferase (Lowy

et al., *Cell* 22, 817-23, 1980) genes which can be employed in *tk* or *aprt* cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler et al., *Proc. Natl. Acad. Sci.* 77, 3567-70, 1980), *npt* confers resistance to the  
5 aminoglycosides, neomycin and G-418 (Colbere-Garapin et al., *J. Mol. Biol.* 150, 1-14, 1981), and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murray, 1992, *supra*). Additional selectable genes have been described. For example, *trpB* allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine  
10 (Hartman & Mulligan, *Proc. Natl. Acad. Sci.* 85, 8047-51, 1988). Visible markers such as anthocyanins,  $\beta$ -glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes et al., *Methods Mol. Biol.* 55, 121-131, 1995).

15

#### Detecting Expression

Although the presence of marker gene expression suggests that the dipeptidyl  
peptidase 8 polynucleotide is also present, its presence and expression may need to  
20 be confirmed. For example, if a sequence encoding a dipeptidyl peptidase 8 polypeptide is inserted within a marker gene sequence, transformed cells containing sequences that encode a dipeptidyl peptidase 8 polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a dipeptidyl peptidase 8 polypeptide under the  
25 control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the dipeptidyl peptidase 8 polynucleotide.

Alternatively, host cells which contain a dipeptidyl peptidase 8 polynucleotide and which express a dipeptidyl peptidase 8 polypeptide can be identified by a variety of  
30 procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or

immunoassay techniques that include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein. For example, the presence of a polynucleotide sequence encoding a dipeptidyl peptidase 8 polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a dipeptidyl peptidase 8 polypeptide. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding a dipeptidyl peptidase 8 polypeptide to detect transformants that contain a dipeptidyl peptidase 8 polynucleotide.

A variety of protocols for detecting and measuring the expression of a dipeptidyl peptidase 8 polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a dipeptidyl peptidase 8 polypeptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton *et al.*, SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med.* 158, 1211-1216, 1983).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding dipeptidyl peptidase 8 polypeptides include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding a dipeptidyl peptidase 8 polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially

available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

5

Expression and Purification of Polypeptides

Host cells transformed with nucleotide sequences encoding a dipeptidyl peptidase 8 polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode dipeptidyl peptidase 8 polypeptides can be designed to contain signal sequences which direct secretion of soluble dipeptidyl peptidase 8 polypeptides through a prokaryotic or eukaryotic cell membrane or which direct the membrane insertion of membrane-bound dipeptidyl peptidase 8 polypeptide.

As discussed above, other constructions can be used to join a sequence encoding a dipeptidyl peptidase 8 polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). Inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the dipeptidyl peptidase 8 polypeptide also can be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a dipeptidyl peptidase 8 polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilized metal ion affinity chromatography, as described

in Porath *et al.*, *Prot. Exp. Purif.* 3, 263-281, 1992), while the enterokinase cleavage site provides a means for purifying the dipeptidyl peptidase 8 polypeptide from the fusion protein. Vectors that contain fusion proteins are disclosed in Kroll *et al.*, *DNA Cell Biol.* 12, 441-453, 1993.

5

#### Chemical Synthesis

Sequences encoding a dipeptidyl peptidase 8 polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers *et al.*,  
10 *Nucl. Acids Res. Symp. Ser.* 215-223, 1980; Horn *et al. Nucl. Acids Res. Symp. Ser.* 225-232, 1980). Alternatively, a dipeptidyl peptidase 8 polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques (Merrifield, *J. Am. Chem. Soc.* 85, 2149-2154, 1963; Roberge *et al.*, *Science* 269, 202-204, 1995). Protein synthesis  
15 can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Optionally, fragments of dipeptidyl peptidase 8 polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

20

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, *PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES*, WH Freeman and Co., New York, N.Y., 1983). The composition of a synthetic dipeptidyl peptidase 8 polypeptide can be confirmed by  
25 amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure; see Creighton, *supra*). Additionally, any portion of the amino acid sequence of the dipeptidyl peptidase 8 polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

30

### Production of Altered Polypeptides

As will be understood by those of skill in the art, it may be advantageous to produce dipeptidyl peptidase 8 polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life that is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter dipeptidyl peptidase 8 polypeptide-encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the polypeptide or mRNA product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

### Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of a dipeptidyl peptidase 8 polypeptide. "Antibody" as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, F(ab')<sub>2</sub>, and Fv, which are capable of binding an epitope of a dipeptidyl peptidase 8 polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, *e.g.*, at least 15, 25, or 50 amino acids.



An antibody which specifically binds to an epitope of a dipeptidyl peptidase 8 polypeptide can be used therapeutically, as well as in immunochemical assays, such as Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody that specifically binds to the immunogen.

10

Typically, an antibody which specifically binds to a dipeptidyl peptidase 8 polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to dipeptidyl peptidase 8 polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a dipeptidyl peptidase 8 polypeptide from solution.

15

Human dipeptidyl peptidase 8 polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, a dipeptidyl peptidase 8 polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (e.g., aluminum hydroxide), and surface active substances (e.g. lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used in humans, BCG (*bacilli Calmette-Guerin*) and *Corynebacterium parvum* are especially useful.

20

25

Monoclonal antibodies that specifically bind to a dipeptidyl peptidase 8 polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not

30

limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler *et al.*, *Nature* 256, 495-497, 1985; Kozbor *et al.*, *J. Immunol. Methods* 81, 31-42, 1985; Cote *et al.*, *Proc. Natl. Acad. Sci.* 80, 2026-2030, 1983; Cole *et al.*, *Mol. Cell Biol.* 62, 109-120, 1984).

5

In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison *et al.*, *Proc. Natl. Acad. Sci.* 81, 6851-6855, 1984; Neuberger *et al.*, *Nature* 312, 604-608, 1984; Takeda *et al.*, *Nature* 314, 452-454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, humanized antibodies can be produced using recombinant methods, as described in GB2188638B. Antibodies that specifically bind to a dipeptidyl peptidase 8 polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. 5,565,332.

15

20

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies that specifically bind to dipeptidyl peptidase 8 polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, *Proc. Natl. Acad. Sci.* 88, 11120-23, 1991).

25

30

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J.*

*Cancer Prev.* 5, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol.* 15, 159-63. Construction of bivalent, bispecific single-chain antibodies is taught in  
5 Mallender & Voss, 1994, *J. Biol. Chem.* 269, 199-206.

A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding  
10 sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology (Verhaar *et al.*, 1995, *Int. J. Cancer* 61, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth.* 165, 81-91).

Antibodies which specifically bind to dipeptidyl peptidase 8 polypeptides also can be  
15 produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi *et al.*, *Proc. Natl. Acad. Sci.* 86, 3833-3837, 1989; Winter *et al.*, *Nature* 349, 293-299, 1991).

20 Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in WO 93/03151. Binding proteins which are derived from immunoglobulins and which are multivalent and multispecific, such as the "diabodies" described in WO 94/13804, also can be prepared.

25 Antibodies according to the invention can be purified by methods well known in the art. For example, antibodies can be affinity purified by passage over a column to which a dipeptidyl peptidase 8 polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

Antisense Oligonucleotides

Antisense oligonucleotides are nucleotide sequences that are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary  
5 nucleotides combine with natural sequences produced by the cell to form complexes and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 11 nucleotides in length, but can be at least 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a  
10 cell as described above to decrease the level of dipeptidyl peptidase 8 gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an automated  
15 synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such as alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters. See Brown, *Meth. Mol. Biol.* 20, 1-8, 1994; Sonveaux, *Meth. Mol. Biol.* 26, 1-72, 1994; Uhlmann *et al.*, *Chem. Rev.* 90, 543-583, 1990.

Modifications of dipeptidyl peptidase 8 gene expression can be obtained by designing antisense oligonucleotides that will form duplexes to the control, 5', or regulatory  
25 regions of the dipeptidyl peptidase 8 gene. Oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription  
30 factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature (*e.g.*, Gee *et al.*, in Huber & Carr, MOLECULAR AND IMMUNOLOGIC

APPROACHES, Futura Publishing Co., Mt. Kisco, N.Y., 1994). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

5     Precise complementarity is not required for successful complex formation between  
an antisense oligonucleotide and the complementary sequence of a dipeptidyl pepti-  
dase 8 polynucleotide. Antisense oligonucleotides which comprise, for example, 2,  
3, 4, or 5 or more stretches of contiguous nucleotides which are precisely comple-  
mentary to a dipeptidyl peptidase 8 polynucleotide, each separated by a stretch of  
10    contiguous nucleotides which are not complementary to adjacent dipeptidyl peptidase  
8 nucleotides, can provide sufficient targeting specificity for dipeptidyl peptidase 8  
mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least  
4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening  
sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can  
15    easily use the calculated melting point of an antisense-sense pair to determine the  
degree of mismatching which will be tolerated between a particular antisense  
oligonucleotide and a particular dipeptidyl peptidase 8 polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to  
20    hybridize to a dipeptidyl peptidase 8 polynucleotide. These modifications can be  
internal or at one or both ends of the antisense molecule. For example, internu-  
cleoside phosphate linkages can be modified by adding cholesteryl or diamine  
moieties with varying numbers of carbon residues between the amino groups and  
terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or  
25    a 3', 5'-substituted oligonucleotide in which the 3' hydroxyl group or the 5'  
phosphate group are substituted, also can be employed in a modified antisense  
oligonucleotide. These modified oligonucleotides can be prepared by methods well  
known in the art. See, e.g., Agrawal *et al.*, *Trends Biotechnol.* 10, 152-158, 1992;  
Uhlmann *et al.*, *Chem. Rev.* 90, 543-584, 1990; Uhlmann *et al.*, *Tetrahedron. Lett.*  
30    215, 3539-3542, 1987.

Ribozymes

Ribozymes are RNA molecules with catalytic activity. See, e.g., Cech, *Science* 236, 1532-1539; 1987; Cech, *Ann. Rev. Biochem.* 59, 543-568; 1990, Cech, *Curr. Opin. Struct. Biol.* 2, 605-609; 1992, Couture & Stinchcomb, *Trends Genet.* 12, 510-515, 1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (e.g., Haseloff *et al.*, U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The coding sequence of a dipeptidyl peptidase 8 polynucleotide can be used to generate ribozymes that will specifically bind to mRNA transcribed from the dipeptidyl peptidase 8 polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art (see Haseloff *et al.* *Nature* 334, 585-591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.*, EP 321,201).

Specific ribozyme cleavage sites within a dipeptidyl peptidase 8 RNA target can be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. Suitability of candidate dipeptidyl peptidase 8 RNA targets also can be evaluated by testing accessibility to hybridi-

zation with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related such that upon hybridizing to the target RNA  
5 through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or  
10 calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease dipeptidyl peptidase 8 expression. Alternatively, if it is desired that the cells stably retain the DNA construct, the construct can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. A  
15 ribozyme-encoding DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in Haseloff *et al.*, U.S. Patent 5,641,673, ribozymes can be engineered so  
20 that ribozyme expression will occur in response to factors that induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

### 25 Differentially Expressed Genes

Described herein are methods for the identification of genes whose products interact with human dipeptidyl peptidase 8. Such genes may represent genes that are differentially expressed in disorders including, but not limited to, cancer, CNS  
30 disorders, and COPD. Further, such genes may represent genes that are differentially regulated in response to manipulations relevant to the progression or treatment of

- 35 -

such diseases. Additionally, such genes may have a temporally modulated expression, increased or decreased at different stages of tissue or organism development. A differentially expressed gene may also have its expression modulated under control versus experimental conditions. In addition, the human dipeptidyl peptidase  
5 8 gene or gene product may itself be tested for differential expression.

The degree to which expression differs in a normal versus a diseased state need only be large enough to be visualized via standard characterization techniques such as differential display techniques. Other such standard characterization techniques by  
10 which expression differences may be visualized include but are not limited to, quantitative RT (reverse transcriptase), PCR, and Northern analysis.

Identification of Differentially Expressed Genes

15 To identify differentially expressed genes total RNA or, preferably, mRNA is isolated from tissues of interest. For example, RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique that does not select against the isolation of mRNA may be utilized for the purification of such RNA samples. See, for example, Ausubel *et*  
20 *al.*, ed., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. New York, 1987-1993. Large numbers of tissue samples may readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, U.S. Patent 4,843,155.

25 Transcripts within the collected RNA samples that represent RNA produced by differentially expressed genes are identified by methods well known to those of skill in the art. They include, for example, differential screening (Tedder *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 85, 208-12, 1988), subtractive hybridization (Hedrick *et al.*,  
*Nature* 308, 149-53; Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 88, 2825, 1984), and,  
30 preferably, differential display (Liang & Pardee, *Science* 257, 967-71, 1992; U.S. Patent 5,262,311).



The differential expression information may itself suggest relevant methods for the treatment of disorders involving the human dipeptidyl peptidase 8. For example, treatment may include a modulation of expression of the differentially expressed  
5 genes and/or the gene encoding the human dipeptidyl peptidase 8. The differential expression information may indicate whether the expression or activity of the differentially expressed gene or gene product or the human dipeptidyl peptidase 8 gene or gene product are up-regulated or down-regulated.

10 Screening Methods

The invention provides assays for screening test compounds that bind to or modulate the activity of a dipeptidyl peptidase 8 polypeptide or a dipeptidyl peptidase 8 polynucleotide. A test compound preferably binds to a dipeptidyl peptidase 8 polypeptide  
15 or polynucleotide. More preferably, a test compound decreases or increases dipeptidyl peptidase activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

Test Compounds

20

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced re-  
25 combinantly, or synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the "one-bead one-compound" library method, and  
30 synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches

are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. See Lam, *Anticancer Drug Des.* 12, 145, 1997.

5 Methods for the synthesis of molecular libraries are well known in the art (see, for example, DeWitt *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 90, 6909, 1993; Erb *et al.* *Proc. Natl. Acad. Sci. U.S.A.* 91, 11422, 1994; Zuckermann *et al.*, *J. Med. Chem.* 37, 2678, 1994; Cho *et al.*, *Science* 261, 1303, 1993; Carell *et al.*, *Angew. Chem. Int. Ed. Engl.* 33, 2059, 1994; Carell *et al.*, *Angew. Chem. Int. Ed. Engl.* 33, 2061; Gallop *et al.*, *J. Med. Chem.* 37, 1233, 1994). Libraries of compounds can be presented in solution  
10 (see, e.g., Houghten, *BioTechniques* 13, 412-421, 1992), or on beads (Lam, *Nature* 354, 82-84, 1991), chips (Fodor, *Nature* 364, 555-556, 1993), bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 89, 1865-1869, 1992), or phage (Scott & Smith, *Science* 249, 386-390, 1990; Devlin, *Science* 249, 404-406, 1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci.* 97, 6378-6382,  
15 1990; Felici, *J. Mol. Biol.* 222, 301-310, 1991; and Ladner, U.S. Patent 5,223,409).

#### High Throughput Screening

20 Test compounds can be screened for the ability to bind to dipeptidyl peptidase 8 polypeptides or polynucleotides or to affect dipeptidyl peptidase 8 activity or dipeptidyl peptidase 8 gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well microtiter plates. The wells of the microtiter plates  
25 typically require assay volumes that range from 50 to 500  $\mu$ l. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the 96-well format.

30 Alternatively, "free format assays," or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by

Jayawickreme *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 19, 1614-18 (1994). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially released the compounds from the beads. Active compounds can be visualized as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa. (Nov. 7-10; 1995). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV-light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described by Salmon *et al.*, *Molecular Diversity* 2, 57-63 (1996). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support. When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

### Binding Assays

For binding assays, the test compound is preferably a small molecule that binds to and occupies, for example, the active site of the dipeptidyl peptidase 8 polypeptide, such that normal biological activity is prevented. Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules.

In binding assays, either the test compound or the dipeptidyl peptidase 8 polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound that is bound to the dipeptidyl peptidase 8 polypeptide can then be accomplished, for example, by direct counting of radioemission, by scintillation counting, or by determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a dipeptidyl peptidase 8 polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a dipeptidyl peptidase 8 polypeptide. A microphysiometer (*e.g.*, Cytosensor™) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a dipeptidyl peptidase 8 polypeptide (McConnell *et al.*, *Science* 257, 1906-1912, 1992).

Determining the ability of a test compound to bind to a dipeptidyl peptidase 8 polypeptide also can be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA) (Sjolander & Urbaniczky, *Anal. Chem.* 63, 2338-2345, 1991, and Szabo *et al.*, *Curr. Opin. Struct. Biol.* 5, 699-705, 1995). BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore™). Changes in the optical phenomenon surface

plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

5 In yet another aspect of the invention, a dipeptidyl peptidase 8 polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent 5,283,317; Zervos *et al.*, *Cell* 72, 223-232, 1993; Madura *et al.*, *J. Biol. Chem.* 268, 12046-12054, 1993; Bartel *et al.*, *BioTechniques* 14, 920-924, 1993; Iwabuchi *et al.*, *Oncogene* 8, 1693-1696, 1993; and Brent W094/10300), to identify other proteins which bind to or interact with the dipeptidyl peptidase 8 polypeptide and  
10 modulate its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct, polynucleotide  
15 encoding a dipeptidyl peptidase 8 polypeptide can be fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct a DNA sequence that encodes an unidentified protein ("prey" or "sample") can be fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to  
20 interact *in vivo* to form an protein-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the  
25 functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein that interacts with the dipeptidyl peptidase 8 polypeptide.

It may be desirable to immobilize either the dipeptidyl peptidase 8 polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound  
30 forms of one or both of the interactants, as well as to accommodate automation of the assay. Thus, either the dipeptidyl peptidase 8 polypeptide (or polynucleotide) or the

test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach the enzyme polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide (or polynucleotide) or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a dipeptidyl peptidase 8 polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, the dipeptidyl peptidase 8 polypeptide is a fusion protein comprising a domain that allows the dipeptidyl peptidase 8 polypeptide to be bound to a solid support. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the non-adsorbed dipeptidyl peptidase 8 polypeptide; the mixture is then incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing proteins or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either a dipeptidyl peptidase 8 polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated dipeptidyl peptidase 8

polypeptides (or polynucleotides) or test compounds can be prepared from biotin-NHS(N-hydroxysuccinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.) and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which  
5 specifically bind to a dipeptidyl peptidase 8 polypeptide, polynucleotide, or a test compound, but which do not interfere with a desired binding site, such as the active site of the dipeptidyl peptidase 8 polypeptide, can be derivatized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

10 Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies which specifically bind to the dipeptidyl peptidase 8 polypeptide or test compound, enzyme-linked assays which rely on detecting an activity of the dipeptidyl peptidase 8 polypeptide, and SDS gel electrophoresis under non-reducing conditions.

15 Screening for test compounds which bind to a dipeptidyl peptidase 8 polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a dipeptidyl peptidase 8 polypeptide or polynucleotide can be used in a cell-based assay system. A dipeptidyl peptidase 8 polynucleotide can be naturally occurring in  
20 the cell or can be introduced using techniques such as those described above. Binding of the test compound to a dipeptidyl peptidase 8 polypeptide or polynucleotide is determined as described above.

#### Enzyme Assays

25 Test compounds can be tested for the ability to increase or decrease the dipeptidyl peptidase activity of a human dipeptidyl peptidase 8 polypeptide. Dipeptidyl peptidase activity can be measured, for example, as described in Maes *et al.*, Neuro-psychopharmacology. 2001 Feb;24(2):130-40; Sentandreu & Toldra, J Agric Food  
30 Chem. 2000 Oct;48(10):5014-22; Li *et al.*, Biochem Biophys Res Commun. 2000 Sep 24;276(2):553-8; or Durinx *et al.*, Eur J Biochem. 2000 Sep;267(17):5608-13.

Enzyme assays can be carried out after contacting either a purified dipeptidyl  
peptidase 8 polypeptide, a cell membrane preparation, or an intact cell with a test  
compound. A test compound that decreases a dipeptidyl peptidase activity of a  
5 dipeptidyl peptidase 8 polypeptide by at least about 10, preferably about 50, more  
preferably about 75, 90, or 100% is identified as a potential therapeutic agent for  
decreasing dipeptidyl peptidase 8 activity. A test compound which increases a  
dipeptidyl peptidase activity of a human dipeptidyl peptidase 8 polypeptide by at  
least about 10, preferably about 50, more preferably about 75, 90, or 100% is  
10 identified as a potential therapeutic agent for increasing human dipeptidyl peptidase 8  
activity.

#### Gene Expression

15 In another embodiment, test compounds that increase or decrease dipeptidyl  
peptidase 8 gene expression are identified. A dipeptidyl peptidase 8 polynucleotide  
is contacted with a test compound, and the expression of an RNA or polypeptide  
product of the dipeptidyl peptidase 8 polynucleotide is determined. The level of  
expression of appropriate mRNA or polypeptide in the presence of the test compound  
20 is compared to the level of expression of mRNA or polypeptide in the absence of the  
test compound. The test compound can then be identified as a modulator of  
expression based on this comparison. For example, when expression of mRNA or  
polypeptide is greater in the presence of the test compound than in its absence, the  
test compound is identified as a stimulator or enhancer of the mRNA or polypeptide  
25 expression. Alternatively, when expression of the mRNA or polypeptide is less in  
the presence of the test compound than in its absence, the test compound is identified  
as an inhibitor of the mRNA or polypeptide expression.

The level of dipeptidyl peptidase 8 mRNA or polypeptide expression in the cells can  
30 be determined by methods well known in the art for detecting mRNA or polypeptide.  
Either qualitative or quantitative methods can be used. The presence of polypeptide



products of a dipeptidyl peptidase 8 polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined *in vivo*, in a cell culture, or in an *in vitro* translation system by detecting incorporation of labeled amino acids into a dipeptidyl peptidase 8 polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell that expresses a dipeptidyl peptidase 8 polynucleotide can be used in a cell-based assay system. The dipeptidyl peptidase 8 polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, such as CHO or human embryonic kidney 293 cells, can be used.

#### Pharmaceutical Compositions

The invention also provides pharmaceutical compositions that can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise, for example, a dipeptidyl peptidase 8 polypeptide, dipeptidyl peptidase 8 polynucleotide, ribozymes or antisense oligonucleotides, antibodies which specifically bind to a dipeptidyl peptidase 8 polypeptide, or mimetics, activators, or inhibitors of a dipeptidyl peptidase 8 polypeptide activity. The compositions can be administered alone or in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries that facilitate processing of the active compounds into preparations which can be

used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Pharmaceutical preparations that can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a

filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

5

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances that increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents that increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

20

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

30

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

#### Therapeutic Indications and Methods

Human dipeptidyl peptidase 8 can be regulated to treat cancer, CNS disorders, and COPD.

#### Cancer

Cancer is a disease fundamentally caused by oncogenic cellular transformation. There are several hallmarks of transformed cells that distinguish them from their normal counterparts and underlie the pathophysiology of cancer. These include uncontrolled cellular proliferation, unresponsiveness to normal death-inducing signals (immortalization), increased cellular motility and invasiveness, increased ability to recruit blood supply through induction of new blood vessel formation (angiogenesis), genetic instability, and dysregulated gene expression. Various combinations of these aberrant physiologies, along with the acquisition of drug-resistance frequently lead to an intractable disease state in which organ failure and patient death ultimately ensue.

Most standard cancer therapies target cellular proliferation and rely on the differential proliferative capacities between transformed and normal cells for their efficacy. This approach is hindered by the facts that several important normal cell types are also highly proliferative and that cancer cells frequently become resistant to these agents. Thus, the therapeutic indices for traditional anti-cancer therapies rarely exceed 2.0.

The advent of genomics-driven molecular target identification has opened up the possibility of identifying new cancer-specific targets for therapeutic intervention that will provide safer, more effective treatments for cancer patients. Thus, newly discovered tumor-associated genes and their products can be tested for their role(s) in disease and used as tools to discover and develop innovative therapies. Genes playing important roles in any of the physiological processes outlined above can be characterized as cancer targets.

Genes or gene fragments identified through genomics can readily be expressed in one or more heterologous expression systems to produce functional recombinant proteins. These proteins are characterized *in vitro* for their biochemical properties and then used as tools in high-throughput molecular screening programs to identify chemical modulators of their biochemical activities. Activators and/or inhibitors of target protein activity can be identified in this manner and subsequently tested in cellular and *in vivo* disease models for anti-cancer activity. Optimization of lead compounds with iterative testing in biological models and detailed pharmacokinetic and toxicological analyses form the basis for drug development and subsequent testing in humans.

#### CNS disorders

Central and peripheral nervous system disorders also can be treated, such as primary and secondary disorders after brain injury, disorders of mood, anxiety disorders, disorders of thought and volition, disorders of sleep and wakefulness, diseases of the motor unit, such as neurogenic and myopathic disorders, neurodegenerative disorders such as Alzheimer's and Parkinson's disease, and processes of peripheral and chronic pain.

Pain that is associated with CNS disorders also can be treated by regulating the activity of human dipeptidyl peptidase 8. Pain which can be treated includes that associated with central nervous system disorders, such as multiple sclerosis, spinal

cord injury, sciatica, failed back surgery syndrome, traumatic brain injury, epilepsy, Parkinson's disease, post-stroke, and vascular lesions in the brain and spinal cord (e.g., infarct, hemorrhage, vascular malformation). Non-central neuropathic pain includes that associated with post mastectomy pain, reflex sympathetic dystrophy (RSD), trigeminal neuralgioradioculopathy, post-surgical pain, HIV/AIDS related pain, cancer pain, metabolic neuropathies (e.g., diabetic neuropathy, vasculitic neuropathy secondary to connective tissue disease), paraneoplastic polyneuropathy associated, for example, with carcinoma of lung, or leukemia, or lymphoma, or carcinoma of prostate, colon or stomach, trigeminal neuralgia, cranial neuralgias, and post-herpetic neuralgia. Pain associated with cancer and cancer treatment also can be treated, as can headache pain (for example, migraine with aura, migraine without aura, and other migraine disorders), episodic and chronic tension-type headache, tension-type like headache, cluster headache, and chronic paroxysmal hemicrania.

15 COPD

Chronic obstructive pulmonary (or airways) disease (COPD) is a condition defined physiologically as airflow obstruction that generally results from a mixture of emphysema and peripheral airway obstruction due to chronic bronchitis (Senior & Shapiro, *Pulmonary Diseases and Disorders*, 3d ed., New York, McGraw-Hill, 1998, pp. 659-681, 1998; Barnes, *Chest* 117, 10S-14S, 2000). Emphysema is characterized by destruction of alveolar walls leading to abnormal enlargement of the air spaces of the lung. Chronic bronchitis is defined clinically as the presence of chronic productive cough for three months in each of two successive years. In COPD, airflow obstruction is usually progressive and is only partially reversible. By far the most important risk factor for development of COPD is cigarette smoking, although the disease does occur in non-smokers.

Chronic inflammation of the airways is a key pathological feature of COPD (Senior & Shapiro, 1998). The inflammatory cell population comprises increased numbers of macrophages, neutrophils, and CD8<sup>+</sup> lymphocytes. Inhaled irritants, such as cigarette

smoke, activate macrophages which are resident in the respiratory tract, as well as epithelial cells leading to release of chemokines (*e.g.*, interleukin-8) and other chemotactic factors. These chemotactic factors act to increase the neutrophil/monocyte trafficking from the blood into the lung tissue and airways. Neutrophils and monocytes recruited into the airways can release a variety of potentially damaging mediators such as proteolytic enzymes and reactive oxygen species. Matrix degradation and emphysema, along with airway wall thickening, surfactant dysfunction, and mucus hypersecretion, all are potential sequelae of this inflammatory response that lead to impaired airflow and gas exchange.

This invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a test compound identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a modulating agent, an antisense nucleic acid molecule, a specific antibody, ribozyme, or a dipeptidyl peptidase 8 polypeptide binding molecule) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

A reagent which affects dipeptidyl peptidase 8 activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce dipeptidyl peptidase 8 activity. The reagent preferably binds to an expression product of a human dipeptidyl peptidase 8 gene. If the expression product is a protein, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells that have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung, liver, spleen, heart brain, lymph nodes, and skin.

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5  $\mu\text{g}$  of DNA per 16 nmole of liposome delivered to about  $10^6$  cells, more preferably about 1.0  $\mu\text{g}$  of DNA per 16 nmole of liposome delivered to about  $10^6$  cells, and even more preferably about 2.0  $\mu\text{g}$  of DNA per 16 nmol of liposome delivered to about  $10^6$  cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition and/or liposomes having a cholesterol backbone conjugated to polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a particular cell type, such as a cell-specific ligand exposed on the outer surface of the liposome.

Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods that are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1  $\mu\text{g}$  to about 10  $\mu\text{g}$  of polynucleotide is combined with about 8 nmol of liposomes, more preferably from



about 0.5  $\mu$ g to about 5  $\mu$ g of polynucleotides are combined with about 8 nmol liposomes, and even more preferably about 1.0  $\mu$ g of polynucleotides is combined with about 8 nmol liposomes.

- 5 In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis *et al.* *Trends in Biotechnol.* 11, 202-05 (1993); Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem.* 263, 621-24 (1988);  
10 Wu *et al.*, *J. Biol. Chem.* 269, 542-46 (1994); Zenke *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 87, 3655-59 (1990); Wu *et al.*, *J. Biol. Chem.* 266, 338-42 (1991).

Determination of a Therapeutically Effective Dose

- 15 The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases dipeptidyl peptidase 8 activity relative to the dipeptidyl peptidase 8 activity which occurs in the absence of the therapeutically effective dose.

- 20 For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses  
25 and routes for administration in humans.

- Therapeutic efficacy and toxicity, *e.g.*, ED<sub>50</sub> (the dose therapeutically effective in 50% of the population) and LD<sub>50</sub> (the dose lethal to 50% of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental  
30 animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD<sub>50</sub>/ED<sub>50</sub>.

Pharmaceutical compositions that exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED<sub>50</sub> with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active ingredient or to maintain the desired effect. Factors that can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either *ex vivo* or *in vivo* using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads,

protoplast fusion, viral infection, electroporation, "gene gun," and DEAE- or calcium phosphate-mediated transfection.

Effective *in vivo* dosages of an antibody are in the range of about 5 µg to about 50 µg/kg, about 50 µg to about 5 mg/kg, about 100 µg to about 500 µg/kg of patient body weight, and about 200 to about 250 µg/kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective *in vivo* dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg, about 1 µg to about 2 mg, about 5 µg to about 500 µg, and about 20 µg to about 100 µg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides that express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of a dipeptidyl peptidase 8 gene or the activity of a dipeptidyl peptidase 8 polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a dipeptidyl peptidase 8 gene or the activity of a dipeptidyl peptidase 8 polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to dipeptidyl peptidase 8-specific mRNA, quantitative RT-PCR, immunologic detection of a dipeptidyl peptidase 8 polypeptide, or measurement of dipeptidyl peptidase 8 activity.

In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act synergistically to effect the treatment or prevention of the various disorders described above. Using this

approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Any of the therapeutic methods described above can be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

#### Diagnostic Methods

Human dipeptidyl peptidase 8 also can be used in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases and abnormalities related to the presence of mutations in the nucleic acid sequences that encode the enzyme. For example, differences can be determined between the cDNA or genomic sequence encoding dipeptidyl peptidase 8 in individuals afflicted with a disease and in normal individuals. If a mutation is observed in some or all of the afflicted individuals but not in normal individuals, then the mutation is likely to be the causative agent of the disease.

Sequence differences between a reference gene and a gene having mutations can be revealed by the direct DNA sequencing method. In addition, cloned DNA segments can be employed as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. For example, a sequencing primer can be used with a double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures using radiolabeled nucleotides or by automatic sequencing procedures using fluorescent tags.

Genetic testing based on DNA sequence differences can be carried out by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized, for example, by high resolution gel electrophoresis. DNA fragments of different se-

quences can be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (*see, e.g., Myers et al., Science 230, 1242, 1985*). Sequence changes at specific locations can also be  
5 revealed by nuclease protection assays, such as RNase and S 1 protection or the chemical cleavage method (*e.g., Cotton et al., Proc. Natl. Acad. Sci. USA 85, 4397-4401, 1985*). Thus, the detection of a specific DNA sequence can be performed by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes and Southern blotting of genomic DNA.  
10 In addition to direct methods such as gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

Altered levels of dipeptidyl peptidase 8 also can be detected in various tissues. Assays used to detect levels of the receptor polypeptides in a body sample, such as  
15 blood or a tissue biopsy, derived from a host are well known to those of skill in the art and include radioimmunoassays, competitive binding assays, Western blot analysis, and ELISA assays.

All patents and patent applications cited in this disclosure are expressly incorporated  
20 herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples, which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

**EXAMPLE 1***Detection of dipeptidyl peptidase 8 activity*

5 The polynucleotide of SEQ ID NO: 1 is inserted into the expression vector pCEV4 and the expression vector pCEV4-dipeptidyl peptidase 8 polypeptide obtained is transfected into human embryonic kidney 293 cells. From these cells extracts are obtained and incubated with substrate in 70 µl phosphate buffer, pH 7.4, for 30 min at 37°C. The specific DPPIV substrates are Gly-Pro-toluenesulfonate, H-Gly-Pro-p-  
10 nitroanilide (NA)/HCl (Sigma, St. Louis, MO, USA) and Gly-Pro-7-amino-4-trifluoromethylcoumarin (Calbiochem, San Diego, CA, USA). It is shown that the polypeptide of SEQ ID NO: 2 hydrolyzes the DPPIV substrates. Thus this polypeptide has a dipeptidyl peptidase activity.

15 **EXAMPLE 2**

*Expression of recombinant human dipeptidyl peptidase 8*

The *Pichia pastoris* expression vector pPICZB (Invitrogen, San Diego, CA) is used  
20 to produce large quantities of recombinant human dipeptidyl peptidase 8 polypeptides in yeast. The dipeptidyl peptidase 8-encoding DNA sequence is derived from SEQ ID NO: 1. Before insertion into vector pPICZB, the DNA sequence is modified by well known methods in such a way that it contains at its 5'-end an initiation codon and at its 3'-end an enterokinase cleavage site, a His6 reporter tag  
25 and a termination codon. Moreover, at both termini recognition sequences for restriction endonucleases are added and after digestion of the multiple cloning site of pPICZ B with the corresponding restriction enzymes the modified DNA sequence is ligated into pPICZB. This expression vector is designed for inducible expression in *Pichia pastoris*, driven by a yeast promoter. The resulting pPICZ/md-His6 vector is  
30 used to transform the yeast.

- 58 -

The yeast is cultivated under usual conditions in 5 liter shake flasks and the recombinantly produced protein isolated from the culture by affinity chromatography (Ni-NTA-Resin) in the presence of 8 M urea. The bound polypeptide is eluted with buffer, pH 3.5, and neutralized. Separation of the polypeptide from the His6 reporter tag is accomplished by site-specific proteolysis using enterokinase (Invitrogen, San Diego, CA) according to manufacturer's instructions. Purified human dipeptidyl peptidase 8 polypeptide is obtained.

### EXAMPLE 3

#### *Identification of test compounds that bind to dipeptidyl peptidase 8 polypeptides*

Purified dipeptidyl peptidase 8 polypeptides comprising a glutathione-S-transferase protein and absorbed onto glutathione-derivatized wells of 96-well microtiter plates are contacted with test compounds from a small molecule library at pH 7.0 in a physiological buffer solution. Human dipeptidyl peptidase 8 polypeptides comprise the amino acid sequence shown in SEQ ID NO: 2. The test compounds comprise a fluorescent tag. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to a dipeptidyl peptidase 8 polypeptide is detected by fluorescence measurements of the contents of the wells. A test compound that increases the fluorescence in a well by at least 15% relative to fluorescence of a well in which a test compound is not incubated is identified as a compound which binds to a dipeptidyl peptidase 8 polypeptide.

**EXAMPLE 4**

*Identification of a test compound which decreases dipeptidyl peptidase 8 gene expression*

5

A test compound is administered to a culture of human cells transfected with a dipeptidyl peptidase 8 expression construct and incubated at 37°C for 10 to 45 minutes. A culture of the same type of cells that have not been transfected is incubated for the same time without the test compound to provide a negative control.

10

RNA is isolated from the two cultures as described in Chirgwin *et al.*, *Biochem.* 18, 5294-99, 1979). Northern blots are prepared using 20 to 30 µg total RNA and hybridized with a <sup>32</sup>P-labeled dipeptidyl peptidase 8-specific probe at 65°C in Express-hyb (CLONTECH). The probe comprises at least 11 contiguous nucleotides selected from the complement of SEQ ID NO: 1. A test compound that decreases the dipeptidyl peptidase 8-specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of dipeptidyl peptidase 8 gene expression.

15

20

**EXAMPLE 5**

*Identification of a test compound which decreases dipeptidyl peptidase 8 activity*

A test compound is administered to a culture of human cells transfected with a dipeptidyl peptidase 8 expression construct and incubated at 37°C for 10 to 45 minutes. A culture of the same type of cells that have not been transfected is incubated for the same time without the test compound to provide a negative control. Dipeptidyl peptidase activity is measured using the method of Maes *et al.*, *Neuropsychopharmacology*. 2001 Feb;24(2):130-40; Sentandreu & Toldra, *J Agric Food Chem*. 2000 Oct;48(10):5014-22; Li *et al.*, *Biochem Biophys Res Commun*.

25

30



2000 Sep 24;276(2):553-8; or Durinx *et al.*, Eur J Biochem. 2000 Sep;267(17):5608-13.

5 A test compound which decreases the dipeptidyl peptidase activity of the dipeptidyl peptidase 8 relative to the dipeptidyl peptidase activity in the absence of the test compound is identified as an inhibitor of dipeptidyl peptidase 8 activity.

#### **EXAMPLE 6**

##### 10 *Tissue-specific expression of dipeptidyl peptidase 8*

The qualitative expression pattern of dipeptidyl peptidase 8 in various tissues is determined by Reverse Transcription-Polymerase Chain Reaction (RT-PCR).

15 To demonstrate that dipeptidyl peptidase 8 is involved in the disease process of COPD, the initial expression panel consists of RNA samples from respiratory tissues and inflammatory cells relevant to COPD: lung (adult and fetal), trachea, freshly isolated alveolar type II cells, cultured human bronchial epithelial cells, cultured small airway epithelial cells, cultured bronchial smooth muscle cells, cultured H441  
20 cells (Clara-like), freshly isolated neutrophils and monocytes, and cultured monocytes (macrophage-like). Body map profiling also is carried out, using total RNA panels purchased from Clontech. The tissues are adrenal gland, bone marrow, brain, colon, heart, kidney, liver, lung, mammary gland, pancreas, prostate, salivary gland, skeletal muscle, small intestine, spleen, stomach, testis, thymus, trachea, thyroid, and  
25 uterus.

To demonstrate that dipeptidyl peptidase 8 is involved in CNS disorders, the following tissues are screened: fetal and adult brain, muscle, heart, lung, kidney, liver, thymus, testis, colon, placenta, trachea, pancreas, kidney, gastric mucosa,  
30 colon, liver, cerebellum, skin, cortex (Alzheimer's and normal), hypothalamus,

cortex, amygdala, cerebellum, hippocampus, choroid, plexus, thalamus, and spinal cord.

To demonstrate that dipeptidyl peptidase 8 is involved in cancer, expression is  
5 determined in the following tissues: adrenal gland, bone marrow, brain, cerebellum,  
colon, fetal brain, fetal liver, heart, kidney, liver, lung, mammary gland, pancreas,  
placenta, prostate, salivary gland, skeletal muscle, small intestine, spinal cord, spleen,  
stomach, testis, thymus, thyroid, trachea, uterus, and peripheral blood lymphocytes.  
Expression in the following cancer cell lines also is determined: DU-145 (prostate),  
10 NCI-H125 (lung), HT-29 (colon), COLO-205 (colon), A-549 (lung), NCI-H460  
(lung), HT-116 (colon), DLD-1 (colon), MDA-MD-231 (breast), LS174T (colon),  
ZF-75 (breast), MDA-MN-435 (breast), HT-1080, MCF-7 (breast), and U87.  
Matched pairs of malignant and normal tissue from the same patient also are tested.

15 *Quantitative expression profiling.* Quantitative expression profiling is performed by  
the form of quantitative PCR analysis called "kinetic analysis" firstly described in  
Higuchi *et al.*, *BioTechnology* 10, 413-17, 1992, and Higuchi *et al.*, *BioTechnology*  
11, 1026-30, 1993. The principle is that at any given cycle within the exponential  
phase of PCR, the amount of product is proportional to the initial number of template  
20 copies.

If the amplification is performed in the presence of an internally quenched  
fluorescent oligonucleotide (TaqMan probe) complementary to the target sequence,  
the probe is cleaved by the 5'-3' endonuclease activity of Taq DNA polymerase and a  
25 fluorescent dye released in the medium (Holland *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*  
88, 7276-80, 1991). Because the fluorescence emission will increase in direct  
proportion to the amount of the specific amplified product, the exponential growth  
phase of PCR product can be detected and used to determine the initial template  
concentration (Heid *et al.*, *Genome Res.* 6, 986-94, 1996, and Gibson *et al.*, *Genome*  
30 *Res.* 6, 995-1001, 1996).

- 62 -

The amplification of an endogenous control can be performed to standardize the amount of sample RNA added to a reaction. In this kind of experiment, the control of choice is the 18S ribosomal RNA. Because reporter dyes with differing emission spectra are available, the target and the endogenous control can be independently quantified in the same tube if probes labeled with different dyes are used.

All "real time PCR" measurements of fluorescence are made in the ABI Prism 7700.

*RNA extraction and cDNA preparation.* Total RNA from the tissues listed above are used for expression quantification. RNAs labeled "from autopsy" were extracted from autaptic tissues with the TRIzol reagent (Life Technologies, MD) according to the manufacturer's protocol.

Fifty  $\mu$ g of each RNA were treated with DNase I for 1 hour at 37°C in the following reaction mix: 0.2 U/ $\mu$ l RNase-free DNase I (Roche Diagnostics, Germany); 0.4 U/ $\mu$ l RNase inhibitor (PE Applied Biosystems, CA); 10 mM Tris-HCl pH 7.9; 10 mM  $MgCl_2$ ; 50 mM NaCl; and 1 mM DTT.

After incubation, RNA is extracted once with 1 volume of phenol:chloroform:isoamyl alcohol (24:24:1) and once with chloroform, and precipitated with 1/10 volume of 3 M NaAcetate, pH5.2, and 2 volumes of ethanol.

Fifty  $\mu$ g of each RNA from the autaptic tissues are DNase treated with the DNA-free kit purchased from Ambion (Ambion, TX). After resuspension and spectrophotometric quantification, each sample is reverse transcribed with the TaqMan Reverse Transcription Reagents (PE Applied Biosystems, CA) according to the manufacturer's protocol. The final concentration of RNA in the reaction mix is 200 ng/ $\mu$ L. Reverse transcription is carried out with 2.5 $\mu$ M of random hexamer primers.

30

*TaqMan quantitative analysis.* Specific primers and probe are designed according to the recommendations of PE Applied Biosystems; the probe can be labeled at the 5' end FAM (6-carboxy-fluorescein) and at the 3' end with TAMRA (6-carboxy-tetramethyl-rhodamine). Quantification experiments are performed on 10 ng of reverse transcribed RNA from each sample. Each determination is done in triplicate.

Total cDNA content is normalized with the simultaneous quantification (multiplex PCR) of the 18S ribosomal RNA using the Pre-Developed TaqMan Assay Reagents (PDAR) Control Kit (PE Applied Biosystems, CA).

The assay reaction mix is as follows: 1X final TaqMan Universal PCR Master Mix (from 2X stock) (PE Applied Biosystems, CA); 1X PDAR control – 18S RNA (from 20X stock); 300 nM forward primer; 900 nM reverse primer; 200 nM probe; 10 ng cDNA; and water to 25 µl.

Each of the following steps are carried out once: pre PCR, 2 minutes at 50°C, and 10 minutes at 95°C. The following steps are carried out 40 times: denaturation, 15 seconds at 95°C, annealing/extension, 1 minute at 60°C.

The experiment is performed on an ABI Prism 7700 Sequence Detector (PE Applied Biosystems, CA). At the end of the run, fluorescence data acquired during PCR are processed as described in the ABI Prism 7700 user's manual in order to achieve better background subtraction as well as signal linearity with the starting target quantity.

#### EXAMPLE 7

*Proliferation inhibition assay: Antisense oligonucleotides suppress the growth of cancer cell lines*

The cell line used for testing is the human colon cancer cell line HCT116. Cells are cultured in RPMI-1640 with 10-15% fetal calf serum at a concentration of 10,000

- 64 -

cells per milliliter in a volume of 0.5 ml and kept at 37°C in a 95% air/5%CO<sub>2</sub> atmosphere.

Phosphorothioate oligoribonucleotides are synthesized on an Applied Biosystems Model 380B DNA synthesizer using phosphoroamidite chemistry. A sequence of 24 bases complementary to the nucleotides at position 1 to 24 of SEQ ID NO: 1 is used as the test oligonucleotide. As a control, another (random) sequence is used: 5'-TCA ACT GAC TAG ATG TAC ATG GAC-3'. Following assembly and deprotection, oligonucleotides are ethanol-precipitated twice, dried, and suspended in phosphate buffered saline at the desired concentration. Purity of the oligonucleotides is tested by capillary gel electrophoresis and ion exchange HPLC. The purified oligonucleotides are added to the culture medium at a concentration of 10 µM once per day for seven days.

The addition of the test oligonucleotide for seven days results in significantly reduced expression of human dipeptidyl peptidase 8 as determined by Western blotting. This effect is not observed with the control oligonucleotide. After 3 to 7 days, the number of cells in the cultures is counted using an automatic cell counter. The number of cells in cultures treated with the test oligonucleotide (expressed as 100%) is compared with the number of cells in cultures treated with the control oligonucleotide. The number of cells in cultures treated with the test oligonucleotide is not more than 30% of control, indicating that the inhibition of human dipeptidyl peptidase 8 has an anti-proliferative effect on cancer cells.

## **EXAMPLE 8**

### ***In vivo testing of compounds/target validation***

#### **1. Acute Mechanistic Assays**

##### ***1.1. Reduction in Mitogenic Plasma Hormone Levels***

This non-tumor assay measures the ability of a compound to reduce either the endogenous level of a circulating hormone or the level of hormone produced in

response to a biologic stimulus. Rodents are administered test compound (p.o., i.p., i.v., i.m., or s.c.). At a predetermined time after administration of test compound, blood plasma is collected. Plasma is assayed for levels of the hormone of interest. If the normal circulating levels of the hormone are too low and/or variable to provide consistent results, the level of the hormone may be elevated by a pre-treatment with a biologic stimulus (i.e., LHRH may be injected i.m. into mice at a dosage of 30 ng/mouse to induce a burst of testosterone synthesis). The timing of plasma collection would be adjusted to coincide with the peak of the induced hormone response. Compound effects are compared to a vehicle-treated control group. An F-test is preformed to determine if the variance is equal or unequal followed by a Student's t-test. Significance is p value  $\leq 0.05$  compared to the vehicle control group.

## 1.2. *Hollow Fiber Mechanism of Action Assay*

Hollow fibers are prepared with desired cell line(s) and implanted intraperitoneally and/or subcutaneously in rodents. Compounds are administered p.o., i.p., i.v., i.m., or s.c. Fibers are harvested in accordance with specific readout assay protocol, these may include assays for gene expression (bDNA, PCR, or Taqman), or a specific biochemical activity (i.e., cAMP levels. Results are analyzed by Student's t-test or Rank Sum test after the variance between groups is compared by an F-test, with significance at  $p \leq 0.05$  as compared to the vehicle control group.

## 2. *Subacute Functional In Vivo Assays*

### 2.1. *Reduction in Mass of Hormone Dependent Tissues*

This is another non-tumor assay that measures the ability of a compound to reduce the mass of a hormone dependent tissue (i.e., seminal vesicles in males and uteri in females). Rodents are administered test compound (p.o., i.p., i.v., i.m., or s.c.) according to a predetermined schedule and for a predetermined duration (i.e., 1

week). At termination of the study, animals are weighed, the target organ is excised, any fluid is expressed, and the weight of the organ is recorded. Blood plasma may also be collected. Plasma may be assayed for levels of a hormone of interest or for levels of test agent. Organ weights may be directly compared or they may be  
5 normalized for the body weight of the animal. Compound effects are compared to a vehicle-treated control group. An F-test is performed to determine if the variance is equal or unequal followed by a Student's t-test. Significance is  $p \text{ value} \leq 0.05$  compared to the vehicle control group.

## 10 2.2. *Hollow Fiber Proliferation Assay*

Hollow fibers are prepared with desired cell line(s) and implanted intraperitoneally and/or subcutaneously in rodents. Compounds are administered p.o., i.p., i.v., i.m., or s.c. Fibers are harvested in accordance with specific readout assay protocol. Cell  
15 proliferation is determined by measuring a marker of cell number (i.e., MTT or LDH). The cell number and change in cell number from the starting inoculum are analyzed by Student's t-test or Rank Sum test after the variance between groups is compared by an F-test, with significance at  $p \leq 0.05$  as compared to the vehicle control group.

20

## 2.3. *Anti-angiogenesis Models*

### 2.3.1. *Corneal Angiogenesis*

25 Hydron pellets with or without growth factors or cells are implanted into a micro-pocket surgically created in the rodent cornea. Compound administration may be systemic or local (compound mixed with growth factors in the hydron pellet). Corneas are harvested at 7 days post implantation immediately following intracardiac infusion of colloidal carbon and are fixed in 10% formalin. Readout is qualitative  
30 scoring and/or image analysis. Qualitative scores are compared by Rank Sum test. Image analysis data is evaluated by measuring the area of neovascularization (in

pixels) and group averages are compared by Student's t-test (2 tail). Significance is  $p \leq 0.05$  as compared to the growth factor or cells only group.

### 2.3.2. *Matrigel Angiogenesis*

5

Matrigel, containing cells or growth factors, is injected subcutaneously. Compounds are administered p.o., i.p., i.v., i.m., or s.c. Matrigel plugs are harvested at pre-determined time point(s) and prepared for readout. Readout is an ELISA-based assay for hemoglobin concentration and/or histological examination (i.e. vessel count, special staining for endothelial surface markers: CD31, factor-8). Readouts are analyzed by Student's t-test, after the variance between groups is compared by an F-test, with significance determined at  $p \leq 0.05$  as compared to the vehicle control group.

15

## 3. Primary Antitumor Efficacy

### 3.1. *Early Therapy Models*

#### 3.1.1. *Subcutaneous Tumor*

20

Tumor cells or fragments are implanted subcutaneously on Day 0. Vehicle and/or compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule starting at a time, usually on Day 1, prior to the ability to measure the tumor burden. Body weights and tumor measurements are recorded 2-3 times weekly. Mean net body and tumor weights are calculated for each data collection day. Anti-tumor efficacy may be initially determined by comparing the size of treated (T) and control (C) tumors on a given day by a Student's t-test, after the variance between groups is compared by an F-test, with significance determined at  $p \leq 0.05$ . The experiment may also be continued past the end of dosing in which case tumor measurements would continue to be recorded to monitor tumor growth delay. Tumor growth delays are expressed as the difference in the median time for the treated and

30



control groups to attain a predetermined size divided by the median time for the control group to attain that size. Growth delays are compared by generating Kaplan-Meier curves from the times for individual tumors to attain the evaluation size. Significance is  $p \leq 0.05$ .

5

### 3.1.2. *Intraperitoneal/Intracranial Tumor Models*

Tumor cells are injected intraperitoneally or intracranially on Day 0. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule starting on Day 1. Observations of morbidity and/or mortality are recorded twice daily. Body weights are measured and recorded twice weekly. Morbidity/mortality data is expressed in terms of the median time of survival and the number of long-term survivors is indicated separately. Survival times are used to generate Kaplan-Meier curves. Significance is  $p \leq 0.05$  by a log-rank test compared to the control group in the experiment.

15

### 3.2. *Established Disease Model*

Tumor cells or fragments are implanted subcutaneously and grown to the desired size for treatment to begin. Once at the predetermined size range, mice are randomized into treatment groups. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Tumor and body weights are measured and recorded 2-3 times weekly. Mean tumor weights of all groups over days post inoculation are graphed for comparison. An F-test is performed to determine if the variance is equal or unequal followed by a Student's t-test to compare tumor sizes in the treated and control groups at the end of treatment. Significance is  $p \leq 0.05$  as compared to the control group. Tumor measurements may be recorded after dosing has stopped to monitor tumor growth delay. Tumor growth delays are expressed as the difference in the median time for the treated and control groups to attain a predetermined size divided by the median time for the control group to attain that size. Growth delays are compared by generating Kaplan-Meier curves from the times

25

30

for individual tumors to attain the evaluation size. Significance is  $p \text{ value} \leq 0.05$  compared to the vehicle control group.

### 3.3. *Orthotopic Disease Models*

5

#### 3.3.1. *Mammary Fat Pad Assay*

10 Tumor cells or fragments, of mammary adenocarcinoma origin, are implanted directly into a surgically exposed and reflected mammary fat pad in rodents. The fat pad is placed back in its original position and the surgical site is closed. Hormones may also be administered to the rodents to support the growth of the tumors. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Tumor and body weights are measured and recorded 2-3 times weekly. Mean tumor weights of all groups over days post inoculation are graphed for  
15 comparison. An F-test is preformed to determine if the variance is equal or unequal followed by a Student's t-test to compare tumor sizes in the treated and control groups at the end of treatment. Significance is  $p \leq 0.05$  as compared to the control group.

20 Tumor measurements may be recorded after dosing has stopped to monitor tumor growth delay. Tumor growth delays are expressed as the difference in the median time for the treated and control groups to attain a predetermined size divided by the median time for the control group to attain that size. Growth delays are compared by generating Kaplan-Meier curves from the times for individual tumors to attain the  
25 evaluation size. Significance is  $p \text{ value} \leq 0.05$  compared to the vehicle control group. In addition, this model provides an opportunity to increase the rate of spontaneous metastasis of this type of tumor. Metastasis can be assessed at termination of the study by counting the number of visible foci per target organ, or measuring the target organ weight. The means of these endpoints are compared by Student's t-test after  
30 conducting an F-test, with significance determined at  $p \leq 0.05$  compared to the control group in the experiment.

### 3.3.2. *Intraprostatic Assay*

5 Tumor cells or fragments, of prostatic adenocarcinoma origin, are implanted directly into a surgically exposed dorsal lobe of the prostate in rodents. The prostate is externalized through an abdominal incision so that the tumor can be implanted specifically in the dorsal lobe while verifying that the implant does not enter the seminal vesicles. The successfully inoculated prostate is replaced in the abdomen and the incisions through the abdomen and skin are closed. Hormones may also be  
10 administered to the rodents to support the growth of the tumors. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Body weights are measured and recorded 2-3 times weekly. At a predetermined time, the experiment is terminated and the animal is dissected. The size of the primary tumor is measured in three dimensions using either a caliper or an ocular micrometer  
15 attached to a dissecting scope. An F-test is performed to determine if the variance is equal or unequal followed by a Student's t-test to compare tumor sizes in the treated and control groups at the end of treatment. Significance is  $p \leq 0.05$  as compared to the control group. This model provides an opportunity to increase the rate of spontaneous metastasis of this type of tumor. Metastasis can be assessed at  
20 termination of the study by counting the number of visible foci per target organ (i.e., the lungs), or measuring the target organ weight (i.e., the regional lymph nodes). The means of these endpoints are compared by Student's t-test after conducting an F-test, with significance determined at  $p \leq 0.05$  compared to the control group in the experiment.

25

### 3.3.3. *Intrabronchial Assay*

Tumor cells of pulmonary origin may be implanted intrabronchially by making an incision through the skin and exposing the trachea. The trachea is pierced with the  
30 beveled end of a 25 gauge needle and the tumor cells are inoculated into the main bronchus using a flat-ended 27 gauge needle with a 90° bend. Compounds are

administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Body weights are measured and recorded 2-3 times weekly. At a predetermined time, the experiment is terminated and the animal is dissected. The size of the primary tumor is measured in three dimensions using either a caliper or an ocular micrometer attached to a dissecting scope. An F-test is preformed to determine if the variance is equal or unequal followed by a Student's t-test to compare tumor sizes in the treated and control groups at the end of treatment. Significance is  $p \leq 0.05$  as compared to the control group. This model provides an opportunity to increase the rate of spontaneous metastasis of this type of tumor. Metastasis can be assessed at termination of the study by counting the number of visible foci per target organ (i.e., the contralateral lung), or measuring the target organ weight. The means of these endpoints are compared by Student's t-test after conducting an F-test, with significance determined at  $p \leq 0.05$  compared to the control group in the experiment.

#### 3.3.4. *Intracecal Assay.*

Tumor cells of gastrointestinal origin may be implanted intracecally by making an abdominal incision through the skin and externalizing the intestine. Tumor cells are inoculated into the cecal wall without penetrating the lumen of the intestine using a 27 or 30 gauge needle. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Body weights are measured and recorded 2-3 times weekly. At a predetermined time, the experiment is terminated and the animal is dissected. The size of the primary tumor is measured in three dimensions using either a caliper or an ocular micrometer attached to a dissecting scope. An F-test is preformed to determine if the variance is equal or unequal followed by a Student's t-test to compare tumor sizes in the treated and control groups at the end of treatment. Significance is  $p \leq 0.05$  as compared to the control group. This model provides an opportunity to increase the rate of spontaneous metastasis of this type of tumor. Metastasis can be assessed at termination of the study by counting the number of visible foci per target organ (i.e., the liver), or measuring the target organ weight. The means of these endpoints are compared by Student's t-test after conducting an F-test,

with significance determined at  $p \leq 0.05$  compared to the control group in the experiment.

#### 4. Secondary (Metastatic) Antitumor Efficacy

##### 4.1. *Spontaneous Metastasis*

Tumor cells are inoculated s.c. and the tumors allowed to grow to a predetermined range for spontaneous metastasis studies to the lung or liver. These primary tumors are then excised. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule which may include the period leading up to the excision of the primary tumor to evaluate therapies directed at inhibiting the early stages of tumor metastasis. Observations of morbidity and/or mortality are recorded daily. Body weights are measured and recorded twice weekly. Potential endpoints include survival time, numbers of visible foci per target organ, or target organ weight. When survival time is used as the endpoint the other values are not determined. Survival data is used to generate Kaplan-Meier curves. Significance is  $p \leq 0.05$  by a log-rank test compared to the control group in the experiment. The mean number of visible tumor foci, as determined under a dissecting microscope, and the mean target organ weights are compared by Student's t-test after conducting an F-test, with significance determined at  $p \leq 0.05$  compared to the control group in the experiment for both of these endpoints.

##### 4.2. *Forced Metastasis*

Tumor cells are injected into the tail vein, portal vein, or the left ventricle of the heart in experimental (forced) lung, liver, and bone metastasis studies, respectively. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Observations of morbidity and/or mortality are recorded daily. Body weights are measured and recorded twice weekly. Potential endpoints include survival time, numbers of visible foci per target organ, or target organ weight. When

survival time is used as the endpoint the other values are not determined. Survival data is used to generate Kaplan-Meier curves. Significance is  $p \leq 0.05$  by a log-rank test compared to the control group in the experiment. The mean number of visible tumor foci, as determined under a dissecting microscope, and the mean target organ weights are compared by Student's t-test after conducting an F-test, with significance at  $p \leq 0.05$  compared to the vehicle control group in the experiment for both endpoints.

### EXAMPLE 9

#### *In vivo testing of compounds/target validation*

##### 1. Pain

##### *Acute Pain*

Acute pain is measured on a hot plate mainly in rats. Two variants of hot plate testing are used: In the classical variant animals are put on a hot surface (52 to 56°C) and the latency time is measured until the animals show nocifensive behavior, such as stepping or foot licking. The other variant is an increasing temperature hot plate where the experimental animals are put on a surface of neutral temperature. Subsequently this surface is slowly but constantly heated until the animals begin to lick a hind paw. The temperature which is reached when hind paw licking begins is a measure for pain threshold.

Compounds are tested against a vehicle treated control group. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

### *Persistent Pain*

Persistent pain is measured with the formalin or capsaicin test, mainly in rats. A solution of 1 to 5% formalin or 10 to 100 µg capsaicin is injected into one  
5 hind paw of the experimental animal. After formalin or capsaicin application the animals show nocifensive reactions like flinching, licking and biting of the affected paw. The number of nocifensive reactions within a time frame of up to 90 minutes is a measure for intensity of pain.

10 Compounds are tested against a vehicle treated control group. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to formalin or capsaicin administration.

### *Neuropathic Pain*

15 Neuropathic pain is induced by different variants of unilateral sciatic nerve injury mainly in rats. The operation is performed under anesthesia. The first variant of sciatic nerve injury is produced by placing loosely constrictive  
20 ligatures around the common sciatic nerve. The second variant is the tight ligation of about the half of the diameter of the common sciatic nerve. In the next variant, a group of models is used in which tight ligations or transections are made of either the L5 and L6 spinal nerves, or the L<sub>6</sub> spinal nerve only. The fourth variant involves an axotomy of two of the three terminal branches  
25 of the sciatic nerve (tibial and common peroneal nerves) leaving the remaining sural nerve intact whereas the last variant comprises the axotomy of only the tibial branch leaving the sural and common nerves uninjured. Control animals are treated with a sham operation.

30 Postoperatively, the nerve injured animals develop a chronic mechanical allodynia, cold allodynia, as well as a thermal hyperalgesia. Mechanical

allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA; Electronic von Frey System, Somedic Sales AB, Hörby, Sweden). Thermal hyperalgesia is measured by means of a radiant heat source (Plantar Test, Ugo Basile, Comerio, Italy), or by means of a cold plate of 5 to 10°C where the nocifensive reactions of the affected hind paw are counted as a measure of pain intensity. A further test for cold induced pain is the counting of nocifensive reactions, or duration of nocifensive responses after plantar administration of acetone to the affected hind limb. Chronic pain in general is assessed by registering the circadian rhythms in activity (Surjo and Arndt, Universität zu Köln, Cologne, Germany), and by scoring differences in gait (foot print patterns; FOOTPRINTS program, Klapdor et al., 1997. A low cost method to analyze footprint patterns. J. Neurosci. Methods 75, 49-54).

Compounds are tested against sham operated and vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

#### *Inflammatory Pain*

Inflammatory pain is induced mainly in rats by injection of 0.75 mg carrageenan or complete Freund's adjuvant into one hind paw. The animals develop an edema with mechanical allodynia as well as thermal hyperalgesia. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA). Thermal hyperalgesia is measured by means of a radiant heat source (Plantar Test, Ugo Basile, Comerio, Italy, Paw thermal stimulator, G. Ozaki, University of California, USA). For edema measurement two methods are being used. In the first method, the animals are sacrificed and the affected hindpaws sectioned and weighed. The second



method comprises differences in paw volume by measuring water displacement in a plethysmometer (Ugo Basile, Comerio, Italy).

5       Compounds are tested against uninflamed as well as vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

### *Diabetic Neuropathic Pain*

10

Rats treated with a single intraperitoneal injection of 50 to 80 mg/kg streptozotocin develop a profound hyperglycemia and mechanical allodynia within 1 to 3 weeks. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA).

15

Compounds are tested against diabetic and non-diabetic vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

20

## **2.     Parkinson's disease**

### *6-Hydroxydopamine (6-OH-DA) Lesion*

25

Degeneration of the dopaminergic nigrostriatal and striatopallidal pathways is the central pathological event in Parkinson's disease. This disorder has been mimicked experimentally in rats using single/sequential unilateral stereotaxic injections of 6-OH-DA into the medium forebrain bundle (MFB).

30

Male Wistar rats (Harlan Winkelmann, Germany), weighing  $200 \pm 250$  g at the beginning of the experiment, are used. The rats are maintained in a temperature- and humidity-controlled environment under a 12 h light/dark cycle with free access to food and water when not in experimental sessions.

5 The following in vivo protocols are approved by the governmental authorities. All efforts are made to minimize animal suffering, to reduce the number of animals used, and to utilize alternatives to in vivo techniques.

10 Animals are administered pargyline on the day of surgery (Sigma, St. Louis, MO, USA; 50 mg/kg i.p.) in order to inhibit metabolism of 6-OHDA by monoamine oxidase and desmethylinipramine HCl (Sigma; 25 mg/kg i.p.) in order to prevent uptake of 6-OHDA by noradrenergic terminals. Thirty minutes later the rats are anesthetized with sodium pentobarbital (50 mg/kg) and placed in a stereotaxic frame. In order to lesion the DA nigrostriatal

15 pathway 4  $\mu$ l of 0.01% ascorbic acid-saline containing 8  $\mu$ g of 6-OHDA HBr (Sigma) are injected into the left medial fore-brain bundle at a rate of 1  $\mu$ l/min (2.4 mm anterior, 1.49 mm lateral, -2.7 mm ventral to Bregma and the skull surface). The needle is left in place an additional 5 min to allow diffusion to occur.

20

#### Stepping Test

Forelimb akinesia is assessed three weeks following lesion placement using a modified stepping test protocol. In brief, the animals are held by the

25 experimenter with one hand fixing the hindlimbs and slightly raising the hind part above the surface. One paw is touching the table, and is then moved slowly sideways (5 s for 1 m), first in the forehand and then in the backhand direction. The number of adjusting steps is counted for both paws in the backhand and forehand direction of movement. The sequence of testing is

30 right paw forehand and backhand adjusting stepping, followed by left paw forehand and backhand directions. The test is repeated three times on three

consecutive days, after an initial training period of three days prior to the first testing. Forehand adjusted stepping reveals no consistent differences between lesioned and healthy control animals. Analysis is therefore restricted to backhand adjusted stepping.

5

#### Balance Test

Balance adjustments following postural challenge are also measured during the stepping test sessions. The rats are held in the same position as described in the stepping test and, instead of being moved sideways, tilted by the experimenter towards the side of the paw touching the table. This maneuver results in loss of balance and the ability of the rats to regain balance by forelimb movements is scored on a scale ranging from 0 to 3. Score 0 is given for a normal forelimb placement. When the forelimb movement is delayed but recovery of postural balance detected, score 1 is given. Score 2 represents a clear, yet insufficient, forelimb reaction, as evidenced by muscle contraction, but lack of success in recovering balance, and score 3 is given for no reaction of movement. The test is repeated three times a day on each side for three consecutive days after an initial training period of three days prior to the first testing.

10

15

20

#### Staircase Test (Paw Reaching)

A modified version of the staircase test is used for evaluation of paw reaching behavior three weeks following primary and secondary lesion placement. Plexiglass test boxes with a central platform and a removable staircase on each side are used. The apparatus is designed such that only the paw on the same side at each staircase can be used, thus providing a measure of independent forelimb use. For each test the animals are left in the test boxes for 15 min. The double staircase is filled with 7 x 3 chow pellets (Precision food pellets, formula: P, purified rodent diet, size 45 mg; Sandown Scientific)

25

30

on each side. After each test the number of pellets eaten (successfully retrieved pellets) and the number of pellets taken (touched but dropped) for each paw and the success rate (pellets eaten/pellets taken) are counted separately. After three days of food deprivation (12 g per animal per day) the  
5 animals are tested for 11 days. Full analysis is conducted only for the last five days.

#### *MPTP treatment*

10 The neurotoxin 1-methyl-4-phenyl-1,2,3,6-tetrahydro-pyridine (MPTP) causes degeneration of mesencephalic dopaminergic (DAergic) neurons in rodents, non-human primates, and humans and, in so doing, reproduces many of the symptoms of Parkinson's disease. MPTP leads to a marked decrease in the levels of dopamine and its metabolites, and in the number of dopaminergic  
15 terminals in the striatum as well as severe loss of the tyrosine hydroxylase (TH)-immunoreactive cell bodies in the substantia nigra, pars compacta.

In order to obtain severe and long-lasting lesions, and to reduce mortality, animals receive single injections of MPTP, and are then tested for severity of  
20 lesion 7–10 days later. Successive MPTP injections are administered on days 1, 2 and 3. Animals receive application of 4 mg/kg MPTP hydrochloride (Sigma) in saline once daily. All injections are intraperitoneal (i.p.) and the MPTP stock solution is frozen between injections. Animals are decapitated on  
day 11.

25

#### Immunohistology

At the completion of behavioral experiments, all animals are anaesthetized with 3 ml thiopental (1 g/40 ml i.p., Tyrol Pharma). The mice are perfused  
30 transcardially with 0.01 M PBS (pH 7.4) for 2 min, followed by 4% paraformaldehyde (Merck) in PBS for 15 min. The brains are removed and placed in

4% paraformaldehyde for 24 h at 4°C. For dehydration they are then transferred to a 20% sucrose (Merck) solution in 0.1 M PBS at 4°C until they sink. The brains are frozen in methylbutan at -20°C for 2 min and stored at -70°C. Using a sledge microtome (mod. 3800-Frigocut, Leica), 25 µm sections are  
5 taken from the genu of the corpus callosum (AP 1.7 mm) to the hippocampus (AP 21.8 mm) and from AP 24.16 to AP 26.72. Forty-six sections are cut and stored in assorters in 0.25 M Tris buffer (pH 7.4) for immunohistochemistry.

A series of sections is processed for free-floating tyrosine hydroxylase (TH) immunohistochemistry. Following three rinses in 0.1 M PBS, endogenous  
10 peroxidase activity is quenched for 10 min in 0.3% H<sub>2</sub>O<sub>2</sub> ±PBS. After rinsing in PBS, sections are preincubated in 10% normal bovine serum (Sigma) for 5 min as blocking agent and transferred to either primary anti-rat TH rabbit antiserum (dilution 1:2000).

Following overnight incubation at room temperature, sections for TH immunoreactivity are rinsed in PBS (2 x10 min) and incubated in biotinylated anti-rabbit immunoglobulin G raised in goat (dilution 1:200) (Vector) for 90 min, rinsed repeatedly and transferred to Vectastain ABC (Vector) solution  
20 for 1 h. 3,3' -Diaminobenzidine tetrahydrochloride (DAB; Sigma) in 0.1 M PBS, supplemented with 0.005% H<sub>2</sub>O<sub>2</sub>, serves as chromogen in the subsequent visualization reaction. Sections are mounted on to gelatin-coated slides, left to dry overnight, counter-stained with hematoxylin dehydrated in ascending alcohol concentrations and cleared in butylacetate. Coverslips are  
25 mounted on entellan.

#### Rotarod Test

We use a modification of the procedure described by Rozas and Labandeira-Garcia (1997), with a CR-1 Rotamex system (Columbus Instruments, Columbus, OH) comprising an IBM-compatible personal computer, a CIO-24  
30

data acquisition card, a control unit, and a four-lane rotarod unit. The rotarod unit consists of a rotating spindle (diameter 7.3 cm) and individual compartments for each mouse. The system software allows preprogramming of session protocols with varying rotational speeds (0–80 rpm). Infrared beams are used to detect when a mouse has fallen onto the base grid beneath the rotarod. The system logs the fall as the end of the experiment for that mouse, and the total time on the rotarod, as well as the time of the fall and all the set-up parameters, are recorded. The system also allows a weak current to be passed through the base grid, to aid training.

### 3. Dementia

#### The object recognition task

The object recognition task has been designed to assess the effects of experimental manipulations on the cognitive performance of rodents. A rat is placed in an open field, in which two identical objects are present. The rats inspect both objects during the first trial of the object recognition task. In a second trial, after a retention interval of for example 24 hours, one of the two objects used in the first trial, the 'familiar' object, and a novel object are placed in the open field. The inspection time at each of the objects is registered. The basic measures in the OR task is the time spent by a rat exploring the two object the second trial. Good retention is reflected by higher exploration times towards the novel than the 'familiar' object.

Administration of the putative cognition enhancer prior to the first trial predominantly allows assessment of the effects on acquisition, and eventually on consolidation processes. Administration of the testing compound after the first trial allows to assess the effects on consolidation processes, whereas administration before the second trial allows to measure effects on retrieval processes.

*The passive avoidance task*

5 The passive avoidance task assesses memory performance in rats and mice. The inhibitory avoidance apparatus consists of a two-compartment box with a light compartment and a dark compartment. The two compartments are separated by a guillotine door that can be operated by the experimenter. A threshold of 2 cm separates the two compartments when the guillotine door is raised. When the door is open, the illumination in the dark compartment is about 2 lux. The light intensity is about 500 lux at the center of the floor of the light compartment.

15 Two habituation sessions, one shock session, and a retention session are given, separated by inter-session intervals of 24 hours. In the habituation sessions and the retention session the rat is allowed to explore the apparatus for 300 sec. The rat is placed in the light compartment, facing the wall opposite to the guillotine door. After an accommodation period of 15 sec. the guillotine door is opened so that all parts of the apparatus can be visited freely. Rats normally avoid brightly lit areas and will enter the dark compartment within a few seconds.

20 In the shock session the guillotine door between the compartments is lowered as soon as the rat has entered the dark compartment with its four paws, and a scrambled 1 mA footshock is administered for 2 sec. The rat is removed from the apparatus and put back into its home cage. The procedure during the retention session is identical to that of the habituation sessions.

30 The step-through latency, that is the first latency of entering the dark compartment (in sec.) during the retention session is an index of the memory performance of the animal; the longer the latency to enter the dark compartment, the better the retention is. A testing compound is given half an

hour before the shock session, together with 1 mg\*kg<sup>-1</sup> scopolamine. Scopolamine impairs the memory performance during the retention session 24 hours later. If the test compound increases the enter latency compared with the scopolamine-treated controls, is likely to possess cognition enhancing potential.

### *The Morris water escape task*

The Morris water escape task measures spatial orientation learning in rodents. It is a test system that has extensively been used to investigate the effects of putative therapeutic on the cognitive functions of rats and mice. The performance of an animal is assessed in a circular water tank with an escape platform that is submerged about 1 cm below the surface of the water. The escape platform is not visible for an animal swimming in the water tank. Abundant extra-maze cues are provided by the furniture in the room, including desks, computer equipment, a second water tank, the presence of the experimenter, and by a radio on a shelf that is playing softly.

The animals receive four trials during five daily acquisition sessions. A trial is started by placing an animal into the pool, facing the wall of the tank. Each of four starting positions in the quadrants north, east, south, and west is used once in a series of four trials; their order is randomized. The escape platform is always in the same position. A trial is terminated as soon as the animal had climbs onto the escape platform or when 90 seconds have elapsed, whichever event occurs first. The animal is allowed to stay on the platform for 30 seconds. Then it is taken from the platform and the next trial is started. If an animal did not find the platform within 90 seconds it is put on the platform by the experimenter and is allowed to stay there for 30 seconds. After the fourth trial of the fifth daily session, an additional trial is given as a probe trial: the platform is removed, and the time the animal spends in the four quadrants is measured for 30 or 60 seconds. In the probe trial, all animals start from the



same start position, opposite to the quadrant where the escape platform had been positioned during acquisition.

5 Four different measures are taken to evaluate the performance of an animal during acquisition training: escape latency, traveled distance, distance to platform, and swimming speed. The following measures are evaluated for the probe trial: time (s) in quadrants and traveled distance (cm) in the four quadrants. The probe trial provides additional information about how well an animal learned the position of the escape platform. If an animal spends more  
10 time and swims a longer distance in the quadrant where the platform had been positioned during the acquisition sessions than in any other quadrant, one concludes that the platform position has been learned well.

In order to assess the effects of putative cognition enhancing compounds, rats  
15 or mice with specific brain lesions which impair cognitive functions, or animals treated with compounds such as scopolamine or MK-801, which interfere with normal learning, or aged animals which suffer from cognitive deficits, are used.

#### 20 *The T-maze spontaneous alternation task*

The T-maze spontaneous alternation task (TeMCAT) assesses the spatial memory performance in mice. The start arm and the two goal arms of the T-maze are provided with guillotine doors which can be operated manually by  
25 the experimenter. A mouse is put into the start arm at the beginning of training. The guillotine door is closed. In the first trial, the 'forced trial', either the left or right goal arm is blocked by lowering the guillotine door. After the mouse has been released from the start arm, it will negotiate the maze, eventually enter the open goal arm, and return to the start position,  
30 where it will be confined for 5 seconds, by lowering the guillotine door. Then, the animal can choose freely between the left and right goal arm (all

guillotine-doors opened) during 14 'free choice' trials. As soon as the mouse has entered one goal arm, the other one is closed. The mouse eventually returns to the start arm and is free to visit whichever goal arm it wants after having been confined to the start arm for 5 seconds. After completion of 14 free choice trials in one session, the animal is removed from the maze. During training, the animal is never handled.

The percent alternations out of 14 trials is calculated. This percentage and the total time needed to complete the first forced trial and the subsequent 14 free choice trials (in s) is analyzed. Cognitive deficits are usually induced by an injection of scopolamine, 30 min before the start of the training session. Scopolamine reduced the per-cent alternations to chance level, or below. A cognition enhancer, which is always administered before the training session, will at least partially, antagonize the scopolamine-induced reduction in the spontaneous alternation rate.

#### REFERENCES

1. Abbott CA, Yu DM, Woollatt E, Sutherland GR, McCaughan GW, Gorrell MD. Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8. Eur J Biochem 2000 Oct;267(20):6140-50
2. Wilson MJ, Ruhland AR, Quast BJ, Reddy PK, Ewing SL, Sinha AA. Dipeptidylpeptidase IV activities are elevated in prostate cancers and adjacent benign hyperplastic glands. J Androl 2000 Mar-Apr;21(2):220-6
3. Raikhlin NT, Bukaeva IA, Probatova NA, Smirnova EA, Pavlovskaya AI, Tupitsyn NN, Sholokhova EN. Dipeptidyl aminopeptidase-IV as a histochemical marker of differential diagnosis of large cell anaplastic CD30+ lymphoma and Hodgkin's disease. Arkh Patol 2000 Jul-Aug;62(4):3-8

4. Hildebrandt M, Reutter W, Arck P, Rose M, Klapp BF. A guardian angel: the involvement of dipeptidyl peptidase IV in psychoneuroendocrine function, nutrition and immune defense. Clin Sci (Colch) 2000 Aug 1;99(2):93-104
5. Umeki K, Yamamoto I, Maruta J, Aratake Y, Ono I, Uemura Y, Tanaka T, Toyoda K, Kotani T, Noguchi S, Sakamoto F, Konoe K, Ohtaki S. CD26/dipeptidyl peptidase IV and thyroid peroxidase as molecular markers for differentiated thyroid carcinoma. Rinsho Byori 1996 Jan;44(1):42-50
10. 6. Van Der Velden VH, Naber BA, Van Hal PT, Overbeek SE, Hoogsteden HC, Versnel MA. Peptidase activities in serum and bronchoalveolar lavage fluid from allergic asthmatics--comparison with healthy non-smokers and smokers and effects of inhaled glucocorticoids. Clin Exp Allergy 1999 Jun;29(6):813-23
15. 23

CLAIMS

1. An isolated polynucleotide being selected from the group consisting of:
  - 5 a) a polynucleotide encoding a dipeptidyl peptidase 8 polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 62% identical to  
the amino acid sequence shown in SEQ ID NO: 2; and  
10 the amino acid sequence shown in SEQ ID NO: 2.
  - b) a polynucleotide comprising the sequence of SEQ ID NO: 1;
  - c) a polynucleotide which hybridizes under stringent conditions to a  
15 polynucleotide specified in (a) and (b) and encodes a dipeptidyl peptidase 8 polypeptide;
  - d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of  
20 the genetic code and encodes a dipeptidyl peptidase 8 polypeptide; and
  - e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d) and encodes a dipeptidyl peptidase 8 polypeptide.  
25
2. An expression vector containing any polynucleotide of claim 1.
3. A host cell containing the expression vector of claim 2.
- 30 4. A substantially purified dipeptidyl peptidase 8 polypeptide encoded by a polynucleotide of claim 1.

5. A method for producing a dipeptidyl peptidase 8 polypeptide, wherein the method comprises the following steps:
  - 5 a) culturing the host cell of claim 3 under conditions suitable for the expression of the dipeptidyl peptidase 8 polypeptide; and
  - b) recovering the dipeptidyl peptidase 8 polypeptide from the host cell culture.
- 10 6. A method for detection of a polynucleotide encoding a dipeptidyl peptidase 8 polypeptide in a biological sample comprising the following steps:
  - 15 a) hybridizing any polynucleotide of claim 1 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and
  - b) detecting said hybridization complex.
- 20 7. The method of claim 6, wherein before hybridization, the nucleic acid material of the biological sample is amplified.
8. A method for the detection of a polynucleotide of claim 1 or a dipeptidyl peptidase 8 polypeptide of claim 4 comprising the steps of:
  - 25 contacting a biological sample with a reagent which specifically interacts with the polynucleotide or the dipeptidyl peptidase 8 polypeptide.
9. A diagnostic kit for conducting the method of any one of claims 6 to 8.
- 30 10. A method of screening for agents which decrease the activity of a dipeptidyl peptidase 8, comprising the steps of:

contacting a test compound with any dipeptidyl peptidase 8 polypeptide encoded by any polynucleotide of claim 1;

5 detecting binding of the test compound to the dipeptidyl peptidase 8 polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential therapeutic agent for decreasing the activity of a dipeptidyl peptidase 8.

10 11. A method of screening for agents which regulate the activity of a dipeptidyl peptidase 8, comprising the steps of:

contacting a test compound with a dipeptidyl peptidase 8 polypeptide encoded by any polynucleotide of claim 1; and

15 detecting a dipeptidyl peptidase 8 activity of the polypeptide, wherein a test compound which increases the dipeptidyl peptidase 8 activity is identified as a potential therapeutic agent for increasing the activity of the dipeptidyl peptidase 8, and wherein a test compound which decreases the dipeptidyl peptidase 8 activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity of the dipeptidyl peptidase 8.

20 12. A method of screening for agents which decrease the activity of a dipeptidyl peptidase 8, comprising the steps of:

25 contacting a test compound with any polynucleotide of claim 1 and detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of dipeptidyl peptidase 8.

30

13. A method of reducing the activity of dipeptidyl peptidase 8, comprising the steps of:
- 5           contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any dipeptidyl peptidase 8 polypeptide of claim 4, whereby the activity of dipeptidyl peptidase 8 is reduced.
14. A reagent that modulates the activity of a dipeptidyl peptidase 8 polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claim 10 to 12.
- 10           15. A pharmaceutical composition, comprising:
- the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.
- 15           16. Use of the expression vector of claim 2 or the reagent of claim 14 in the preparation of a medicament for modulating the activity of a dipeptidyl peptidase 8 in a disease.
- 20           17. Use of claim 16 wherein the disease is , cancer, a CNS disorder or COPD.
18. A cDNA encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NO: 2.
- 25           19. The cDNA of claim 18 which comprises SEQ ID NO: 1.
20. The cDNA of claim 18 which consists of SEQ ID NO: 1.
- 30           21. An expression vector comprising a polynucleotide which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO: 2.

22. The expression vector of claim 21 wherein the polynucleotide consists of SEQ ID NO: 1.
- 5 23. A host cell comprising an expression vector which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NO: 2.
24. The host cell of claim 23 wherein the polynucleotide consists of SEQ ID NO: 1.
- 10 25. A purified polypeptide comprising the amino acid sequence shown in SEQ ID NO: 2.
26. The purified polypeptide of claim 25 which consists of the amino acid sequence shown in SEQ ID NO: 2.
- 15 27. A fusion protein comprising a polypeptide having the amino acid sequence shown in SEQ ID NO: 2.
- 20 28. A method of producing a polypeptide comprising the amino acid sequence shown in SEQ ID NO: 2, comprising the steps of:
- culturing a host cell comprising an expression vector which encodes the polypeptide under conditions whereby the polypeptide is expressed; and
- 25 isolating the polypeptide.
29. The method of claim 28 wherein the expression vector comprises SEQ ID NO: 1.



30. A method of detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NO: 2, comprising the steps of:
- 5 hybridizing a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO: 1 to nucleic acid material of a biological sample, thereby forming a hybridization complex; and
- detecting the hybridization complex.
- 10 31. The method of claim 30 further comprising the step of amplifying the nucleic acid material before the step of hybridizing.
32. A kit for detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NO: 2, comprising:
- 15 a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO: 1; and instructions for the method of claim 30.
33. A method of detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NO: 2, comprising the steps of:
- 20 contacting a biological sample with a reagent that specifically binds to the polypeptide to form a reagent-polypeptide complex; and
- 25 detecting the reagent-polypeptide complex.
34. The method of claim 33 wherein the reagent is an antibody.
35. A kit for detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NO: 2, comprising:
- 30

an antibody which specifically binds to the polypeptide; and

instructions for the method of claim 33.

- 5      36. A method of screening for agents which can modulate the activity of a human dipeptidyl peptidase 8, comprising the steps of:
- 10      contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 62% identical to the amino acid sequence shown in SEQ ID NO: 2 and (2) the amino acid sequence shown in SEQ ID NO: 2; and
- 15      detecting binding of the test compound to the polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential agent for regulating activity of the human dipeptidyl peptidase 8.
- 20      37. The method of claim 36 wherein the step of contacting is in a cell.
- 20      38. The method of claim 36 wherein the cell is *in vitro*.
- 20      39. The method of claim 36 wherein the step of contacting is in a cell-free system.
- 25      40. The method of claim 36 wherein the polypeptide comprises a detectable label.
- 25      41. The method of claim 36 wherein the test compound comprises a detectable label.
- 30      42. The method of claim 36 wherein the test compound displaces a labeled ligand which is bound to the polypeptide.

43. The method of claim 36 wherein the polypeptide is bound to a solid support.

44. The method of claim 36 wherein the test compound is bound to a solid support.

5

45. A method of screening for agents which modulate an activity of a human dipeptidyl peptidase 8, comprising the steps of:

10 contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 62% identical to the amino acid sequence shown in SEQ ID NO: 2 and (2) the amino acid sequence shown in SEQ ID NO: 2; and

15 detecting an activity of the polypeptide, wherein a test compound which increases the activity of the polypeptide is identified as a potential agent for increasing the activity of the human dipeptidyl peptidase 8, and wherein a test compound which decreases the activity of the polypeptide is identified as a potential agent for decreasing the activity of the human dipeptidyl peptidase 8.

20 46. The method of claim 45 wherein the step of contacting is in a cell.

47. The method of claim 45 wherein the cell is *in vitro*.

25 48. The method of claim 45 wherein the step of contacting is in a cell-free system.

49. A method of screening for agents which modulate an activity of a human dipeptidyl peptidase 8, comprising the steps of:

30 contacting a test compound with a product encoded by a polynucleotide which comprises the nucleotide sequence shown in SEQ ID NO: 1; and

detecting binding of the test compound to the product, wherein a test compound which binds to the product is identified as a potential agent for regulating the activity of the human dipeptidyl peptidase 8.

5

50. The method of claim 49 wherein the product is a polypeptide.

51. The method of claim 49 wherein the product is RNA.

10 52. A method of reducing activity of a human dipeptidyl peptidase 8, comprising the step of:

contacting a cell with a reagent which specifically binds to a product encoded by a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO: 1, whereby the activity of a human dipeptidyl peptidase 8 is reduced.

15

53. The method of claim 52 wherein the product is a polypeptide.

54. The method of claim 53 wherein the reagent is an antibody.

20

55. The method of claim 52 wherein the product is RNA.

56. The method of claim 55 wherein the reagent is an antisense oligonucleotide.

25 57. The method of claim 56 wherein the reagent is a ribozyme.

58. The method of claim 52 wherein the cell is *in vitro*.

59. The method of claim 52 wherein the cell is *in vivo*.

30

60. A pharmaceutical composition, comprising:

a reagent which specifically binds to a polypeptide comprising the amino acid sequence shown in SEQ ID NO: 2; and

5 a pharmaceutically acceptable carrier.

61. The pharmaceutical composition of claim 60 wherein the reagent is an antibody.

10 62. A pharmaceutical composition, comprising:

a reagent which specifically binds to a product of a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO: 1; and

15 a pharmaceutically acceptable carrier.

63. The pharmaceutical composition of claim 62 wherein the reagent is a ribozyme.

20 64. The pharmaceutical composition of claim 62 wherein the reagent is an antisense oligonucleotide.

65. The pharmaceutical composition of claim 62 wherein the reagent is an antibody.

25

66. A pharmaceutical composition, comprising:

an expression vector encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NO: 2; and

30

a pharmaceutically acceptable carrier.

67. The pharmaceutical composition of claim 66 wherein the expression vector comprises SEQ ID NO: 1.
- 5 68. A method of treating a dipeptidyl peptidase 8 dysfunction related disease, wherein the disease is selected from, cancer, a CNS disorder or COPD comprising the step of:
- 10 administering to a patient in need thereof a therapeutically effective dose of a reagent that modulates a function of a human dipeptidyl peptidase 8, whereby symptoms of the dipeptidyl peptidase 8 disfunction related disease are ameliorated.
- 15 69. The method of claim 68 wherein the reagent is identified by the method of claim 36.
70. The method of claim 68 wherein the reagent is identified by the method of claim 45.
- 20 71. The method of claim 68 wherein the reagent is identified by the method of claim 49.

Fig. 1

atgCGgaagg ttaagaaact gCGcctggac aaggagaaca cCGgaagttg gagaagcttc  
 tcgctgaatt cCGagggggc tgagaggatg gccaccaccg ggaccccaac ggccgaccga  
 ggCGacgcag cCGccacaga tgaccCGgcc gccCGcttcc aggtgcagaa gcactCGtgg  
 gacgggctcc ggagcatcat ccacggcagc CGcaagtact cgggCctcat tgtcaacaag  
 gCGcccccag acttccagtt tgtgcagaag acggatgagt ctgggccccca ctcccaccgc  
 ctctactacc tgggaatgcc atatggcagc cgtgagaact ccctCctcta ctccgagatc  
 cccaagaaag tgcggaagga ggccctgctg ctgctgtcct ggaagcagat gctggaccac  
 ttccaggcca cccccacca tgggtgtctac tcccGagagg aggagctact cggggagcgc  
 aagcgcctgg gcgtcttcgg aatcacctct tatgacttcc acagtGagag cggcctcttc  
 ctcttccagg ccagcaatag cctgttccac tgcagggatg gtggcaagaa tggctttatg  
 gtgtccccga tgaagccact ggagatcaag actcagtgtt ctgggcccac catggacccc  
 aaaatctgcc ctgCGaccc tgccttcttc tccttcatca ataacagcga cctgtgggtg  
 gccaacatcg agacaggcga ggagcggcgg ctgaccttct gccaccaagg tttatccaat  
 gtctggatg acccaagtc tgcgggtgtg gccaccttCG tcatacagga agagtctgac  
 cgcttactg ggtactggtg gtgcccaca gcctCctggg aaggttcaga gggcctcaag  
 acgctgcgaa tcctgtatga ggaagtCGat gagtccgagg tggaggatcat tcacgtcccc  
 tctctgCGc tagaagaaag gaagacggac tcgtatcggg accccaggac aggcagcaag  
 aatcccaaga ttgccttgaa actggctgag ttccagactg acagccaggg caagatcgtc  
 tcgaccagg agaaggagct ggtgcagccc ttcagctcgc tgttccgaa ggtggagtac  
 atCGccagg ccgggtggac ccgggatggc aaatacgcct gggccatgtt cctggaccgg  
 cccagcagt ggctccagct cgtCctctc cccCGggcc tgttcatccc gagcacagag  
 aatgaggagc agcggctagc ctctgccaga gctgtcccca ggaatgtcca gccgtatgtg  
 gtgtacgagg aggtcaccaa cgtctggatc aatgttcatg acatcttcta tcccttcccc  
 caatcagagg gagaggacga gctctgcttt ctccGcgcca atgaatgcaa gaccggttc  
 tgccatttgt acaaagtcac cgccgtttta aaatcccagg gctacgattg gagtGagccc  
 ttcagccccg gggaagatga atttaagtgc ccattaagg aagagatggc tctgaccagc  
 ggtgaatggg aggttttggc gaggcacggg tccaagatct gggTcaatga ggagaccaag  
 ctggtgtact tccagggcac caaggacacg ccgctggagc accacctcta cgtggtcagc  
 tatgaggcgg cggcgagat cgtacgcctc accacgcccg gcttctccca tagctgtcc  
 atgagccaga acttCGacat gttcgtcagc cactacagca gcgtgagcac gccgcctgc  
 gtgcacgtct acaagctgag cggccccgac gacgacccc tgcacaagca gcccCGcttc  
 tgggctagca tgatggaggc agccagctgc cccCGgatt atgttCctcc agagatcttc  
 catttccaca cgcgctcGga tgtgCGgctc tacggcatga tctacaagcc ccacgccttg  
 cagccaggga agaagcacc caccgtctc tttgtatatg gaggcccca ggtgcagctg  
 gtgaataact ctttcaaagg catcaagtac ttgCGgctca acacactggc ctccctggg  
 tacgCCgtgg ttgtgattga cggcagggg tctgtcagc gagggcttCG gttcgaaggg  
 gccctgaaaa accaaatggg ccagggtggag atCGaggacc aggtggaggg cctgcagttc  
 gtggccgaga agtatggctt cattgacttg agccgagtCG ccatccatgg ctggctctac  
 gggggcttcc tctcgtcat ggggctaata cacaagccc aggtgttcaa ggtggccatc  
 gCGggtgccc cggtcaccgt ctggatggc tacgacacag ggtacactga gcgctacatg  
 gacgtccctg agaacaacca gcacggctat gaggcgggtt ccgtggctct gcacgtggag  
 aagctgcccc atgagcccaa ccgcttgcct atCctccacg gcttCctgga cgaaaacgtg  
 cactttttcc acacaaactt cctcgtctcc caactgatcc gagcagggaa accttaccag  
 ctccagatct accccaacga gagacacagt attcgtgcc ccgagtcggg cgagcactat  
 gaagtcacgt tgctgcactt tctacaggaa tacctc

- 2/13 -

Fig. 2

MRKVKKLRD KENTGSWSRF SLNSEGAERM ATTGTPTADR GDAAATDDPA ARFQVQKHSW  
 DGLRSIIHGS RKYSGLIVNK APHDFQFVQK TDESGPHSHR LYVLGMPYGS RENSLLYSEI  
 PKKVRKEALL LLSWKQMLDH FQATPHHGVY SREEELLRER KRLGVFGITS YDFHSESGLF  
 LFQASNSLFH CRDGGKNGFM VSPMKPLEIK TQCSGPRMDP KICPADPAFF SFINNSDLWV  
 ANIETGEERR LTFCHQGLSN VLDDPKSAGV ATFVIEEED RFTGYWWCPT ASWEGSEGLK  
 TLRILYEEVD ESEVEVIHVP SPALEERKTD SYRYPRTGSK NPKIALKLAE FQTSQSGKIV  
 STQEKELVQP FSSLFPKVEY IARAGWTRDG KYAWAMFLDR PQOWLQLVLL PPALFIPSTE  
 NEEQRLASAR AVPRNVQPYV VYEEVTNVMI NVHDIFYPPF QSEGEDELCE LANECKTGF  
 CHLYKVAVL KSQGYDWSEP FSPGEDEFKC PIKEEMALTS GEWEVLARHG SKIWNNEETK  
 LVYFQGTKDT PLEHHLVVS YEAAGEIVRL TTPGFSSHCS MSQNFDMFVS HYSSVSTPPC  
 VHVKLSGPD DDPLHKQPRF WASMMEAASC PPDYVPPPIF HFHTRSDVRL YGMIYKPHAL  
 QPGKKHPTVL FVYGGPQVQL VNNSFKGIKY LRLNTLASLG YAVVVIDGRG SCQRGLRFEG  
 ALKNQMGQVE IEDQVEGLQF VAEKYGFIDL SRVAIHGWSY GGFLSLMGLI HKPQVFKVAI  
 AGAPVTWMA YDTGYTERYM DVPENNQHGY EAGSVALHVE KLPNEPNRLI ILHGFLDENV  
 HFFHTNFLVS QLIRAGKPYQ LQIYPNERHS IRCPESGEHY EVTLLHFLQE YL

Fig. 3

MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQL  
 KKLADTRKYHGYMMAPKPHDFMFKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPK  
 TINRAAVLMLSWKPLLDLFQATLDYGMYSREEELLRERKRIGTVGIASDYHQSGTFL  
 FQAGSGIYHVKGQGPQGFTQQPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSDIWI  
 SNIVTREERRLTYVHNELANMEEDARSAGVATFVLQEEFDRYSGYWWCPKAETTPSGGK  
 ILRILYEENDESEVEIIHVTSPLMETRRADSFYRPTGTANPKVTFKMSEIMIDAEGRI  
 IDVIDKELIQPFELFEGVEYIARAGWTPEGKYAWSILLDRSQTRLQIVLISPELFIPV  
 EDDVMBRQRLIESVPDSVTPLIYEETTDIWINIHDIFHVFPQSHEEEIEFIFASECKT  
 GFRHLYKITSILKESKYKRSSGGLPAPSDFKCPIKEEIAITSGBEVLGRHGSNIQVDE  
 VRRLVYFEGTKDSPLEHHLVVSYNPGEVTRLTDRGYSHSCCISQHCDFIFISKYSNQK  
 NPHCVSLYKLSSPEDDPTCKTKEFWATILDSAGPLPDYTPPEIFSFESTTGFTLYGMLY  
 KPHDLQPGKKYPTVLFIYGGPQVQLVNNRFGVKYFRLNTLASLYVVVIDNRGSCHR  
 GLKFEGAFKYKMGQIEIDDQVEGLQYLASRYDFIDLDRVGIHGWSYGGYLSLMALMORS  
 DIFRVAIAGAPVTLWIFYDTGYTERYMGHPDQNEQGYLGSVAMQAEKFPSEPNRLLLL  
 HGFLDENVHFAHTSILLSFLVRAGKPYDLQIYPQERHSIRVPESGEHYELHLLHYLQEN  
 LGSRIAALKVI

Fig. 4

GAGGTTCAAGAGGGCTCAAGACGCTGCGAATCCTGTATGAGGAAGTCGATGAGTCCGAGGT  
 GGAGGTCATTCACGTCCCCTCTCCTGCGCTAGAGAAGAAAGGAAGACGGACTCGTATCGGTA  
 CCCAGGACAGGCAGCAAGAATCCCAAGATTGCCTTGAAACTGGCTGAGTTCAGACTGA  
 CAGCCAGGGCAAGATCGTCTCGACCCAGGAGAAGGAGCTGGTGACGCCCTTCAGCTCGCT  
 GTTCCCGAAGGTGGAGTACATCGCCAGGGCCGGGTGGACCCGGGATGGCAAAATACGCCTG  
 GGCCATGTTCTGGACCGGCCCCAGCAGTGGCTCCAGCTCGTCTCTCCCCCGGCCCT  
 GTTCATCCCGAGCACAGAGAATGAGGAGCAGCGGCTAGCCTCTGCCAGAGCTGTCCCCAG  
 GAATGTCCAGCCGTATGTGGTGTACGAGGAGGTACCAACGTCTGGATCAATGTTTCATGA  
 CATCTTCTATCCCTTCCCCAATCAGAGGGAGAGGACGAGCTCTGCTTTCTCCGCGCCAA  
 TGAATGCAAGACCGGCTTCTGCCATTGTACAAAGTCACCGCGGTTTTAAATCCAGGGC  
 TACGATTGGAG



- 3/13 -

Fig. 5

GGGGAGGTTTTGGCGAGGCACGGCTCCAAGGGCACCAAGGACACGCCGCTGGAGCACCAC  
CTCTACGTGGTCAGCTATGAGGCGGCCGCGAGATCGTACGCCCTACCCACGCCCGGCTTC  
TCCCATAGCTGCTCCATGAGCCAGAACTTCGACATGTTCTGTCAGCCACTACAGCAGCGTG  
AGCACGCCCGCCTGCGTGCACGTCTACAAGCTGAGCGGCCCGGACGACGACCCCTGCAC  
AAGCAGCCCCGCTTCTGGGCTAGCATGATGGAGGCAGCCAGCTGCCCGGATTATGTT  
CCTCCAGAGATCTTCCATTCCACACGCGCTCGGATGTGCGGCTCTACGGCATGATCTAC  
AAGCCCCACGCTTGCAGCCAGGGAAGAAGCACCACCGTCTCTTGTATATGGAGGC  
CCCCAGGTGCAGCTGGTGAATAACTCCTTCAAAGGCATCAAGTACTTGGCGCTCAACACA  
CTGGCCTCCCTGGGCTACGCGTGTTGTGATTGACGGCAGGGCTCCTGTGAGCGAGGGC  
TTCGGTTCGAAGGGGCCCTGAAAAAACAAATGGGCCAGGTGGAGATCGAGGACCAGGTGG  
GAGGGCCTGCAGTTCGCTGAGCGAGAAGTATGGCTTCATCCGAAGTACGCGAGTTGCCA  
TCCAATGGCTGGGTCTACGGGACTTACTCTCGCTCATGGGGTAATCCACAAGCCCAGG  
GGGTGAGGGCAATCGCGAGTGCCCGGGCACGGCTGGATGGCCAGAAACAGCGTACACT  
GAGCGGAAGGGACGGCCCGGAAAAACAGACAGGAAAGGGAGTCCGGGGCCGCGTCAGG  
GAGAAAGTGCCACGAAGCCACAGCGGGGATCACACAGAGGCTGCGGACACGGGGGGTT  
CCACAAAAACGACACAAGGAGACGGGAAACACCAACAACAATACCAGGAACAATGAGGC  
GCGGGGAAACGACACGAGAACAACATGGACGCGACGAACAAAGCAATGCGACAGCAGA  
GACGCGGGACGACCAGGACGCCAGTTTCGACGGTTCGCCCTCGCGCGCGGAGCGACCAGN

Fig. 6

GCTCGGATGTGCGGCTCTACGGCATGATCTACAAGCCCCACGCCTTGACGCCAGAGAAGA  
AGCACCCCACCGTCTTGTATATGGAGGCCCCAGGTGCAGCTGGTGAATAACTCCT  
TCAAAGGCATCAAGTACTTGGCGCTCAACACACTGGCCTCCCTGGGCTACGCCGTGGTTG  
TGATTGACGGCAGGGGCTCCTGTGACGAGGGCTTCGGTTCGAAGGGGCCCTGAAAAACC  
AAATGGGCCAGGTGGAGATCGAGGACCAGGTGGAGGGCTGCAGTTCTGTCGCGGAGAAGT  
ATGGCTTCATCGACACTGAGCCGAGTTGCCATCCATGGCTGGTCTACGGGGCTTCCTC  
TCGCTCATGGGGCTAATCCACAAGCCCCAGGTGTTCAAGGTGGCCATCGCGGGTGCCCCG  
GTCACCGTCTGGATGGCCTACGACACAGGGTACACTGAGCGCTACATGGACGTCCCTGAG  
AACAACCAGCACGGCTATGAGGCGGGTTCCGTGGCCCTGCACGTGGAGAAGCTGCCCAAT  
GAGCCCAACGATTGCTTATCCTCCACGGTTCTTGACGAAAACGTGCACTTTTTCCACAC  
AAATTCCTCGTCTCCCACTGATCCGAGCAGGGAAACCTTACAGCGTTCGGATCAACCCAA  
GAGAGAACAGTTTCGCTGCCGAGTCGGCGACATATGAGTCGTTGATGACTTCTAAGGAAA  
CTCTGACCTGCACGGACGCCATTA

Fig. 7

AAAGCAGAGCTCGTCCTCTCCCTCTGATTGGGGGAAGGGATAGAAGATGTCATGAACATT  
GATCCAGACGTTGGTGACCTCCTCGTACACCACATACGGCTGGACATTCTGGGGACAGC  
TCTGGCAGAGGCTAGCCGCTGCTCCTCATTCTCTGTGCTCGGGATGAACAGGGCCGGGGG  
GAGGAGGACGAGCTGGAGCCAATGCTGGGGCCGGTCCAGGAACATGGCCAGGCGTATTT  
GCCATCCCGGGTCCACCCGGCCCTGGCGATGTACTCCACCTTCGGGAACAGCGAGCTGAA  
GGGCTGCACCAGCTCCTTCTCCTGGGTGAGACGATCTTGCCCTGGCTGTCAGTCTGGAA  
CTCAGCCAGTTTCAAGGCAATCTTGGGATTCTTGCTGCTGCTGCTGGGGTACCGATACGA  
GTCCGTCTTCTTCTTCTAGCGCAGGAGAGGGGACGTGAATGACCTCCACCTCGGACTC  
ATCGACTTCTCATAAGGATTCCGA

- 4/13 -

Fig. 8

ATTACATCTATTTTAAAGGAAAGCAAATATAAACGATCCAGTGGTGGGCTGCCTGCTCCA  
AGTGATTTCAAGTGTCTATCAAAGAGGAGATAGCAATTACCAGTGGTGAATGGGAAGTT  
CTTGGCCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAA  
GGCACCAAAGACTCCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGA  
GAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTCTTGCTGCATCAGTCAGCACTGT  
GACTTCTTTATAAGTAAGTATAGTAACCAGAAGAATCCACACTGTGTGTCCCTTTACAAG  
CTATCAAGTCCTGAAGATGACCCAACCTTGCAAAACAAAGGAATTTTGGGCCACCATTTTG  
GATTCAAGCAGGTCTCTTCTGACTATACTCCTCCAGAAATTTTCTCTTTTGAAAGTACT  
ACTGGATTTACATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAA  
TATCCTACTGTGCTGTTTATATATGGTGGTCCCTCAGGTGCAGTTGGTGAATAATCGGTTT  
AAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGCCTCTCTAGGTTATGTGGTTGTAGTG  
ATAGACAACAGGGGATCCTGTACCGANGGCTTAAATTTGAAGGCGCTTTAAATATAAA  
ATGGGTCAAATAGAAATTGACGATCAGGTGGGA

Fig. 9

AAAGCAGAGCTCGTCCTCTCCCTCTGATTGGGGGAAGGGATAGAAGATGTCATGAACATT  
GATCCAGACGTTGGTGACCTCCTCGTACACCACATACGGCTGGACATTCTGGGGACAGC  
TCTGGCAGAGGCTAGCCGCTGCTCCTCATTCTCTGTGCTCGGGATGAACAGGGCCGGGG  
GAGGAGGACGAGCTGGAGCCACTGCTGGGGCCGGTCCAGGAACATGGCCAGGCGTATTT  
GCCATCCCGGGTCCACCCGGCCTGGCGATGTACTCCACCTTCGGGAACAGCGAGCTGAA  
GGGCTGCACCAGCTCCTTCTCCTGGGTGAGACGATCTTGCCCTGGCTGTGAGTCTGGAA  
CTCAGCCAGTTTCAAGGCAATCTTGGGATACTTGCTGCCTGTCTGGGGTACCGATACGA  
GTCCGTCTTCTTTCTTCTAGCGCAGGAGAGGGGACGTGAATGACCTCCACCTCGGACTC  
ATCGACTTCTCAT

Fig. 10

CGGCTCCAAGGGCACCAAGGACACGCCGCTGGAGCACCACCTCTACGTGGTCAGCTATGA  
GGCGGCCGGCGAGATCGTACGCCTCACCACGCCCGGCTTCTCCCATAGCTGCTCCATGAG  
CCAGAACTTCGACATGTTTCGTAGCCACTACAGCAGCGTGAGCACGCCGCCCTGCGTGCA  
CGTCTACAAGCTGAGCGGCCCGACGACGACCCCTGCACAAGCAGCCCCGCTTCTGGGC  
TAGCATGATGGAGGCAGCCAGCTGCCCCCGGATTATGTTCTCCAGAGATCTTCCATT  
CCACACGCGCTCGGATGTGCGGCTCTACGGCATGATCTACAAGCCCCACGCCCTTGACGCC  
AGGGAAGAAGCACCCACCGTCTCTTTGTATATGGAGGCCCCAGGTGCAGCTGGTGA  
ATAACTCCTTCAAAGGCATCAAGTACTTGGCGCTCAACACACTGGGCTCCTGTGGCTAC  
GCCGTGGTTTGTGATTGACGGCAGGGGCTCCTGTGAGCGAGGGCTTCGGTTCGAAGGGGC  
CTGAAAAACCAAATGGGCCAGTGTGGAGATCGAGGACCTGGTGGAGGGCCTGCAGTTTCGT  
GGCCGAGAAGTATGGCTTCTCGACTGAGCCGAGTTGCCCTCCTTGGCTGTCTACGGGGC  
TCTCTCGCTCATGGGGCTATCCACAGCCCGTGTCCGTGGCTCGCGGTGCCGGCCCTTGG  
TGCTACGACCGGTCTGGGCTCTGATCTGGAAC

- 5/13 -

Fig. 11

CGGGCTGTCTGTCCTCCCGCTCCCGCCACTTCCGGGGTTCGAGTCCCGGGCATGGAGCCGCG  
ACCGTGAGGCGCCGCTGGACCCGGGACGACCTGCCAGTCCGGCCGCGCCCGCCACGTCCC  
GGTCTGTGTCCACGCCTGCAGCTGGAATGGAGGCTCTCTGGACCTTTAGAAAGGCACCC  
CTGCCCTCCTGAGGTGAGCTGAGCGGTTAATGCGGAAGGTTAAGAAACTGCGCCTGGACA  
AGGAGAACACCGGAAGTTGGAGAAGCTTCTCGCTGAATTCGAGGGGGCTGAGAGGATGG  
CCACCACCGGGACCCCAACGGCCGACCGAGGCGACGCAGCCGCCACAGATGACCCGGCCG  
CCCGCTTCCAGGTGCAGAAGCACTCGTGGGACGGGCTCCGGAGCATCATCCACGGCAGCC  
GCAAGTACTCGGGCCTCATTTGTCAACAAGGCGCCCCACGACTTCCAGTTTGTGCAGAAGA  
CGGATGAGTCTGGGCCCCACTCCACCCGCTCTACTACCTGGGAATGCCATATGGCAGCC  
GAGAGAACTCCCTCCTCTACTCTGAGATTCCCAAGAAGGTCCGGAAAGAGGCTCTGCTGC  
TCCTTGTCTGGAAGCAGATGCTGGCATCATTTCCAGGCCACGCCCCACATGGGGGTCTA  
CTCTCGGGAGGAGGAGCTGCTGAGGGAGCGGAAACGCTGGGGGCTCTTCGGCATCACCT  
CCTACGACTTCCACAGCGAGAGTGGCCTCTTCTCTTCCAGGCCACACAGCCTTCTCCAC  
TGGCGCGACGGCGGAAGAACGGCTCATGGGTGTCCCTATGAACCGTGGGAATTACACCC  
AAGGGACAGGGCCCCGAGTGGGAGCGAAAATCTGCCCGCGACCATGACGTATATGCCGT  
CAAACAACAACAGGCACCCGTGGGTGTGCCATCCTAACACGGGACGAAACGAGCGGACC  
CTCCGCCCCAAGGGATCACGCCGAGAGACCCAACTCAGTGGGGGGACACTGCAGAGACAA  
CCAACACGAATCGTACACGTGGCGCACCCCGCGACACAGAAGCAACACAGCAACGAGA  
AGCAGACACAGGGCGCAAGCCGACCTAGACAGAGCAGACCGCAGGACGGTACGAGCACA  
ACATCTGAAGACCGCAGCCACCGCCCGC

Fig. 12

GAGTCAAGTATTTCCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATA  
GACAACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGGCGCCTTTAAATATAAAATG  
GGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGAT  
TTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTATGGAGGATACCTCTCCCTG  
ATGGCATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACT  
CTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTACCCCTGACCAGAAT  
GAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCA  
AATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAATGTCCATTTTGACATACCAGT  
ATATTACTGAGTTTFTTAGTGAGGGCTGGAAAGCCATATGATTACAGATCTATCCTCAG  
GAGAGACACAGCATAAGAGTTCTGAATCGGGAGAACATTATGAAGTGCATCTTTGAC  
TACCTTCAAGAAAACCTTGGATCACGTATTGCTGCTCTAAAAGTGATATAAATTTGACCT  
GTGTAGAACTCTCTGGTATACACTGGTATTTGACCAATGAGGACGTGTAATCGAAGCGA  
AAACCCAGAAATGGTCATCGCCATTGTGTACCTGCATTGTTAGCATTTACTTCTGGAAA  
ATTAATGTTGGTGCCATGCGGGCGCTTACGGGTGGGGGAAATTAATACTTTAACCCCATG  
TGCTAATCAATGTATTTCTCGAAACCAGTCTAGATTTCTAAAAAAAACCGAGATTT  
CTTCAGAGGGCGCAAGGGGGCCACAGGGCTTAAAAGGCTGGAAGAGACCGTCAATGTGC  
CAGTGTGCAAACTCTTTCGTGAGAGAATTGTATGATAGGGAG

- 6/13 -

Fig. 13

GGCTTCTGCCATGTGTACAAAGTCACCGCCGTTTTTAAATCCCAGGGCTACGATGGGAGT  
GAGCCCTTCAGCCCCGGGGAAGATGAATTTAAGTGCCCCATTAAAGGAAGAGATGGCTCTG  
ACCAGCGGTGAATGGGAGGTTTGGCGAGGCACGGTTCAGATCTGGGTCAATGAGGAG  
ACCAAGCTGGTGTACTTCCAGGGCACCAAGGACACGCCGCTGGAGCACCACCTTACGTG  
GTCAGCTATGAGGCGGCCGGCGAGATCGTACGCCCTCAACCAGGCCCGGGTTCTCCCATAG  
CTGGCTCCATGAGCCAGAACTTCGACATGTTCTGTAGCCACTACAGCAGCGTGAGCACGC  
CGCCCTGCGTGACGCTTAACAAGCTGGAGCTGGCCCCGANGAGGAACCCCTGCACAAGC  
AGCCCGGGTTCTGGGCTAGCATGATGGAGGCAGCCAGGTGCCCCGGGGATTATGTTCTC  
CAGAGATCTTCCATTTCCACACGCGCTCGGATGTGCGGTCTACGGCATGATCTAGAAGCC  
CAAGCTTGGAGCCAGGACGAAGACCAACGGCCCTCTTGATTGTTGGGGCCAGGTGCCGTGG  
GGCTAATCCTTCAAGGGCTCAGGACATGCGGGTACACAATGGGCTCCTGGGTACGCCGGG  
AAGGAGATGCGGAGGGGGCCGACAGGAGGGGCAGGGCAAAGGGCCGAGAAAAATTGGGC  
CGGGAACCGGACCTGGAGAGGCAAAATTGGCACAACAGTGGTTAAAGCGAGAAAGGGACA  
CCAAGCGAGGGACAGGAGGACCTACAAGTGGGTGACAACCCGCTAGAGAGAGAAGACGG  
ACCCACGGGATGCAAAGAAACACGAGAACCACAAATCAAAGTACGAGAGCGACGAAAC  
AATGCAAGCGTAAGTTCTTCTCGCTGCCCTCGCTCGCGCCGTCTGTCTGCTGTTGGTG  
GGAGAGGAG

Fig. 14

TGGCGAATTCGGCACGAGGGGGAGGTTTTGGCGAGGCACGGCTCCAAGGGCACCAAGGAC  
ACGCCGCTGGAGCACCCTCTACGTGGTCAGCTATGAGGCGGCCGGCGAGATCGTACGC  
CTACCCACGCCCCGGCTTCTCCCATAGCTGCTCCATGAGCCAGAACTTCGACATGTTCTGT  
AGCCACTACAGCAGCGTGAGCAGCGCCCGCTGCGTGACGCTCTACAAGCTGAGCGGCCCC  
GACGACGACCCCTGCACAAGCAGCCCCGCTTCTGGGCTAGCATGATGGAGGCAGCCAGC  
TGCCCCCGGATTATGTTCTTCCAGAGATCTTCCATTTCCACACGCGCTCGGATGTGCGG  
CTCTACGGCATGATCTACAAGCCCCACGCCCTTGACGCCAGGGAAGAAGCACCCACCGTC  
CTCTTTGTATATGGAGGCCCCC

Fig. 15

CAACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGGCGCCTTTAAATATAAAATGGG  
TCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTT  
CATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCCTATGGAGGATACCTCTCCCTGAT  
GGCATTAAATGCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCT  
GTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTGACCAGAATGA  
ACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTCCCCCTCTGAACAAA  
TCGTTTACTGCTCTTACATGGTTTCTGGATGAGAATGTCCATTTTGACATACACAGTAT  
ATTACTGAGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTACAGATCTATCCTCAGGA  
GAGACACAGCATAAGAGTTCTTGAATCGGGAGAACATTATGAACTGCATCTTTTGCACTA  
CCTTCAAGAAAACCTTGATCACGTATTGCTGCTCTAAAAGTGATATAATTTTGACCTGT  
GTAGAACTCTCTGGTATACACTGGCTATTTAACCAAATGAGGAGGTTAATCAACAGAAA  
ACACAGAATTGATCATCATTTTGATACCTGCCATGTAACATCTACTNCTGAAAATAAA  
TGTGGTGCCATGCAG

- 7/13 -

Fig. 16

GTCTACGGGGGCTTCCTCTCGCTCATGGGGCTAATCCACAAGCCCCAGGTGTTCAAGGTG  
GCCATCGCGGGTGCCCCGGTCAACCGTCTGGATGGCCTACGACACAGGGTACACTGAGCGC  
TACATGGACGTCCCTGAGAACCAACCAGCACGGCTATGAGGCGGGTTCCGTGGCTCTGCAC  
GTGGAGAAGCTGCCCAATGAGCCCAACCGCTTGCTTATCCTCCACGGCTTCCTGGACGAA  
AACGTGCACTTTTTCCACACAACTTCTCGTCTCCCACTGATCCGAGCAGGGAAACCT  
TACCAGTCCAGATCTACCCCAACGAGAGACACAGTATTCTGCTGCCCGAGTCGGGCGAG  
CACTATGAAGTCACGTTGCTGCACTTTCTACAGGAATACCTCTGAGCCTGCCACCGGGA  
GCCGCCACATCACAGCACAAAGTGGCTGCAGCCTCCGCGGGGAACAGGCGGGAGGGACTG  
AGTGGCCCGCGGGCCCCAGTGAGGCACTTTGTCCCGCCAGCGCTGGCCAGCCCCGAGGA  
GCCGCTGCCTTCACCGCCCGACGCCTTTTATCCTTTTTTAAACGCTCTTGGGTTTTATGT  
CCGCTGCTTCTTGTTGCCGAGACAGAGAGATGGTGGTCTCGGGCCAGCCCCCTTCTCTC  
CCCGCTTCTGAGGAGGAGGTCACACGCTGATGGGCACTGGAGAGGCAGAAGAGACTC  
AGAGGAGGGGGCTGCTTCCGCTGGGGCTCCTGTGACCTCTTCACTCCCCCTGGCCGGGCA  
GGCCGCTCCAGGCCCAAGCTGCTTTGTGCTCGTCCCCCGGGCGGTTCCCACTTGTGTTGG  
TTGGTGGGGGGATTTCATTTTAAAGAGGGGGGGGGCGCTGTTCCC

Fig. 17

TGCCCCACAGCCTCCTGGGAAGGTTTCAAGAGGCTCAAGACGCTGCGAATCCTGTATGAG  
GAAGTCGATGAGTCCGAGGTGGAGGTCAATCACGTCCCCTCTCCTGCGCTAGAAGAAAGG  
AAGACGGAATCGCATCGGTACCCAGGACAGGCAGCAAGAATCCCAAGATTGCCTTGAAA  
CTGGCTGAGTTCAGACTGACAGCCAGGGCAAGATCGTCTCGACCCAGAGAAAGAGCTG  
GTGCAGCCCTTCAGCTCGCTGTTCCCGAAGGTGGAGTACATCGCCAGGGCCGGGTGGACC  
CGGGATGGCAAATACGCCTGGGCCATGTTCTGGACCGGCCCCAGCAGTGGCTCCAGCTC  
GTCTTCTCCCCCGCCCTGTTTCATCCCGAGCACAGAGAATGAGGAGCAGCGCTAGCC  
TCTGGCGAGCTGTCNCAGGAATGTCCAGCCGTAT

Fig. 18

GCTCTTACCCGACTTTTGCTCTTTTGGGGTCTTTGCTCTTTTGGGGTCTTTGCTCATGGTG  
TCCCTATGAAACCGCTGGAAATCAAGACCCAGTGTCTCAGGGCCCCGGATGGACCCCAAA  
ATCTGCCCTGCCGACCTGCCTTCTTCTCCTTCATCAATAACAGCGACCTGTGGGTGGCC  
AACATCGAGACAGGCGAGGAGCGGCGGTGACCTTCTGCCACCAAGGTTTATCCAATGTC  
CTGGATGACCCCAAGTCTGCGGGTGTGGCCACCTTCGTATACAGGAAGAGTTCGACCGC  
TTCATGGGTACTGGTGGTGCCCCACAGCCTCCTGGGAAGGTTTCAAGGGCCTCAAGAGC  
CTGCGAATCCTGTATGAGGAAGTCGATGAGTCCGAGGTGGAGGTCAATCACGTCCACTCT  
CTGCGCTAGAAGAAAGGAAGACGGACTCGTA

Fig. 19

CGTCGTGTAATCCGTACAATGAACAAACAGCATATTGAACCACACGTGACTAACGCGATG  
AGTCGACAGCATTTCGTAGGCTCTGCTCACCTTCCCCGGGGCTGAAGGGCTCACTCCAAT  
CGTAGCCCTGGGATTTTAAACGGCGGTGACTTTGTACAAATGGCAGAAGCCGGTCTTGC  
AATTCATTGGCGCGGAGAAAGCAGAGCTCGTCTCTCCCTCTGATTTGGGGGAAGGGATAG  
AAGATGTATGAACATTGATCCAGACGTTGGTGACCTCCTCGTACACCACATACGGCTGG  
ACATTCCTGGGGACAGCTCTGGCAGAGGCTAGCCGCTGCTCCTCATTCCTGTGCTCGGG  
ATGAACAGGGCCGGGGGGAGGAGGACGAGCTGGAGCCACTGCTGGGGCCGGTCCAGGAAC  
ATGGCCAGGCGTATTTGCCATCCCGGGTCCACCCGGCCCTGGCGATGTACTCCACCTTC  
GGGAACAGCGAGCTGAAGGGCTGCACCAGCTCCTTCTCCTGGGTGAGACGATCTTGCCC  
TGGCTGTCACTCTGG

- 8/13 -

Fig. 20

GATCTTGCCCTGGCTGTGCTAGTCTGGAAGTCTAGCCAGTTTCAAGGCAATCTTGGGATTCTT  
GCTGCCTGTCTGGGGTACCGATACGAGTCCGTCTTCTTTCTTCTAGCGCAGGAGAGGG  
GACGTGAATGACCTCCACCTCGGACTCATCGACTTCTTCATACAGGATTTCGACGCTCTT  
GAGGCCCTCTGAACCTTCCCAGGAGGCTGTGGGGCACCACCAGTACCCAGTGAAGCGGT  
GAACTCTTCTGTATGACGAAGGTGGCCACACCCGACAGCTTGGGGTCATCCAGGACATT  
GGATAAACCTTGGTGGCAGAAGGTCTAGCCGCCGCTCCTCGCT

Fig. 21

AACGGGTGATGCGGTTGGTTAAGAACTGCGCCTGGACAAGGAGAACACCGGAAGTTGGA  
GAAGCTTCTCGCTGAATCCGAGGGGGCTGAGAGGATGGCCACCACCGGGACCCCAACGG  
CCGACCGAGGCGACGACGCCGCCACAGATGACCCGGCCGCCCGCTTCCAGGTGACAGAAGC  
ACTCGTGGGACGGGCTCCGGAGCATCATCCACGGCAGCCGCAAGTACTCGGGCCTCATTG  
TCAACAAGGCGCCCCACGACTTCCAGTTTGTGCAGAAGACGGATGAGTCTGGGCCCCACT  
CCAACCGCATCACGGT

Fig. 22

TTAATGCGGAAGGTTAAGAACTGCGCCTGGACAAGGAGAACACCGGAAGTTGGAGAAGC  
TTCTCGCTGAATCCGAGGGGGCTGAGAGGATGGCCACCACCGGGACCCCAACGGCCGAC  
CGAGGCGACGCAGCCGCCACAGATGACCCGGCCGCCCGCTTCCAGGTGACAGAAGCACTCG  
TGGGACGGGCTCCGGAGCATCATCCACGGCAGCCGCAAGTACTCGGGCCTCATTGTCAAC  
ATCGAGACAGGCGAGGAGCGCGGCTGACCTTCTGCCACCAAGGTTTATCCAATGTCTTG  
GATGACCCCAAGTCTGCGGGTGTGGCCACCTTCGTTCATACAGGAAGAGTTTCGACCGCTTC  
ACTGGGTACTGGTGGTGGCCACAGCCTCCTGGGAAGGTTTCAGAGGGCCTCAAGACGCTG  
CGAATCCTGTATGAGGAAGTCGATGAGTCCGAGGTGGAGGTCATTACGTACCTCTCCT  
GCGCTAGAAGAAAGGAAGACGGACTCGTATCGGTACCCCA

Fig. 23

AGGCTGTGGGGCACCACCACTACCCAGTGAAGCGGTGCAACTCTTCTGTATGACGAACG  
TGGCCACACCCGACAGCTTGGGGTCATCCAGGACATTGGATAAACCTTGGTGGCAGAAGG  
TCAGCCGCCGCTCCTCGCTGTCTCGCATGTTGGCCACCCACAGGTGCTGTATTGATG  
AAGGAGAAGAAGGCAGGCTCGGCAGGGCAGATTTTGGGGTCCATCCGGGGCCCTGAGCAC  
TGGGTCTTGATTTCCAGCGGTTTCATAGGGGACACCATGAAGCCGTTCTTGCCGCCGTCG  
CGGCAGTGGAAGAGGCTGTTGCTGGCCTGGAAGAGGAAGTGGCCACTCTCGTGTGGAAG  
TCGTAGGAGGTGATGCCGAAGACCCCGAGCGTTTCCGCTCCCTCAGCAGC

Fig. 24

TTGTGGCTCAGTGCCAATAAACTTTATTTATGAACACAGGTGGTGGGCCGGATTGACC  
TGTGGGCCGTAGGTGGCCAACCCCTGGTTGTGCTAAGGGCCTGGCACAGGGGAGGTGCTC  
AGTAAATCTGCTGCCTGAATAAGCCAACAGTTGGGTGCTCTAAGCCCTGCCAGATCATGC  
AGTCACTGGGGAAGGCTGGCTTCTGCCGATCCCTGTTCTTCTCCCGCAGGGTGTGCTG  
GCTGAGTGGGGGGGCCGACCAATGAACAAACAGCATATTGAACCACACGTGACTAACGCG  
ATGAGTCGACAGCATTTCGTGAGGCTCTGCTCACCTTCCCGGGGCTGAAGGGCTCACTCC  
AATCGTAGCCCTGGGATTTTAAAACGGCGGTGACTTTGTACAAATGGCAGAAGCCGGTCT  
TGCATTCAATTGGCGCGGAGAAAGCAGAGCTCGTCTCTCCCTCTGATTGGGGGAAGGGAT  
AGAAGATGTCATGAACATTGATCCAGACGTTGGTGACCTCCTCCGTACACACATACGGCT  
CGGACATTCTGGGGACAGCTCTCGCAGAGGCTAGCCGCCCTGCTCTCATTCTCTGTGCTCG

- 9/13 -

Fig. 25

AGATCGAGGACCAGGTGGAGGGCCTGCAGTTTCGTGGCCGAGAAGTATGGCTTCATCGACC  
TGAGCCGAGTTGCCATCCATGGCTGGTCTACGGGGGCTTCCTCTCGCTCATGGGGCTAA  
TCCACAAGCCCCAGGTGTTCAAGGTGGCCATCGCGGGTGCCCCGGTCACCGTCTGGATGG  
CCTACGACACAGGGTACACCTGAGCGCTACATGGACTGTCCTGAGAACAACAGCACGGC  
TATGAGGCGGGTTCCGTGGCCCTGCCACGTTGTGAGAATGGCCTGCCCCAAATGAGCCCC  
CAAACCCCGCTTTGCCTTTAATCCCNATCCCCACGTGGCGTTTTCCCTNTGTGTTAACCG  
GAAAAACGTTGCAACACTTATATTTTTCGCCAGCGAAGGCCAAAAACAACGGTTTTCC  
CCAATTACGGTGATTCTTTCCCCCAAGAACCTTGGAAAATTTCCCGGTAATGCCAG  
GGGGGAAAAACCTTACCAAGCTCCAGATCTACCCAACGAGAGACACAGTATTTCGTGC  
CCCGAGTCGTGGCGAGCACTATGAAGTCACGTTGCTGCACTTTTTCCCTTAAACAAGGGGG  
AAAATTTAAACCCCTTCCCTTGGAAAGGCCCTTGGCCCCAA

Fig. 26

GACATAAAACCCAAGAGCGTTTAAAGAAAAGGATAAAAGGCGTCGGGGCGGTGAAGGCAG  
GGCTCCTCGGGGCTGGCCAGCGCTGGGCGGGACAAAGTGCCCTACTGGGGCCCCGCGGGC  
CACTCAGTCCCTCCCGCTGGTTCCCCGCGGAGGCTGCAGCCACTTGTGCTGTGATGTGG  
CGGCTCCCGGTGGGCAGGCTCAGAGGTATTCTGTAGAAAGTGACCAACGTGACTTCAT  
AGTGCTCGCCGACTCGGGGCAGCAATACTGTGTCTCTCGTTGGGGTAGATCTGGAGCT  
GGTAAGGTTTCCCTGCTCGGATCAGTTGGGAGACGAGGAAGTTTGTGTGAAAAAGTGCA  
CGTTTTCGTCCAGGAAGCCGTGGAGGATAAGCAAGCGGTTGGGCTCATTGGGCAGCTTCT  
CACGTGCAGGGCCACGGAACCCGCCTC

Fig. 27

GACATAAAACCCAAGAGCGTTTAAAAAAGGATAAAAGGCGTCGGGGCGGTGAAGGCAGC  
GGCTCCTCGGGGCTGGCCAGCGCTGGGCGGGACAAAGTGCCCTACTGGGGCCCCGCGGGC  
ACTCAGTCCCTCCCGCTGGTTCCCCGCGGAGGCTGCAGCCACTTGTGCTGTGATGTGGC  
GGCTCCCGGTGGGCAGGCTCAGAGGTATTCTGTAGAAAGTGACCAACGTGACTTCATA  
GTGCTCGCCGACTCGGGGCAGCAATACTGTGTCTCTCGTTGGGGTAGATCTGGAGCTG  
GTAAGGTTTCCCTGCTCGGATCAGTTGGGAGACGAGGAAGTTTGTGTGAAAAAGTGAC  
GTTTTCGTCCAGGAAGCCGTGGAGGATAAGCAAGCGGTTGGGCTCATTGGGCAGCTTCTC  
CACGTGCAGGGCCACGGAACCC

Fig. 28

TCTGGTTCGTGCGTGGCATGGTGTGGACTTTGGGTGGCTCGTGGGCTTCTCCACGTGCAGG  
CCACGGAACCCGCTCATAGCCGTGCTGGTTGTTCTCAGTGGACGTCCATGTAGCGCTCA  
GTGTACCCTGTGTGCTAGTGCCATCCAGACGGTGACCGGGCACCCGCGATGGCCACCTT  
GAACACCTGGGGCTTGTGGATTAGCCCCATGAGCGAGAGGAAGCCCCGTTAGGACAGCC  
ATGGATGGCAACTCGGCTCAGGTCGATGAAGCCATACTTCTCGGCCACGAACTGCAGGCC  
CTCCACCTGGTCTCGATCTCCACCATGAG

- 10/13 -

Fig. 29

BLASTP - alignment of 181\_Protein against trembl|AF221634|AF221634\_1  
 gene: "DPP8"; product: "dipeptidyl peptidase 8"; Homo sapiens dipeptidyl  
 peptidase 8 (DPP8) mRNA, complete cds. //:gp|AF221634|11095188 gene:  
 "DPP8";  
 product: "dipeptidyl peptidase 8"; Homo sapiens dipeptidyl peptidase 8  
 (DPP8)  
 mRNA, complete cds.

This hit is scoring at : 0.0 (expectation value)

Alignment length (overlap) : 840

Identities : 61 %

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

Database searched : nrdb\_1\_;

Q: 53 FQVQKHSWDGLRSIIHGSRKYSGLIVNKAPHDFQFVQKTDES GPHSHRLYYLGMPYGSRE  
 F.V:::SW. L::: .:RKY.G:: KAPHDF.FV::D..GPHS.R:YYL.M. :RE  
 H: 35 FYVERYSWSQLKKLLADTRKYHGYMMAKAPHDFMFVKRNDPDGPHSDRIYYLAMSGENRE

NSLLYSEIPKKVRKEALLLLSWKQMLDHFQATPHHGVSREEELLRERKRLGVFGITSYD  
 N:L.YSEIPK...A:L:LSWK.:LD FQAT .:G:YSREEELLRERKR:G..GI.SYD  
 NTLFYSEIPKTINRAAVLMLSWKPLLDLFQATLDYGMYSREEELLRERKRIGTVGIASYD

FHSEGLFLFQASNSLFHCRDGGKNGFMVSPMKPLEIKTQCSGPRMDPKICPADPAFFSF  
 :H. SG.FLQA...:H.:DGG..GF. .P::P .:T.C.. RMDPK:CPADP :.:F  
 YHQSGTFLFQAGSGIYHVKDGGPQGFTQQPLRPNLVETSCPNI RMDPKLCPADPDWIAF

INNSDLWVANIETGEERLTFCHQGLSNVLDDPKSAGVATFVIQEEFDRFTGYWWCPTAS  
 I:::D:W::NI T EERLTL:H. L:N: :D.:SAGVATFV:QEEFDR::GYWWCP.A.  
 IHSNDIWNISNIVTREERLTYVHNELANMEEDARSAGVATFVLQEEFDRYSGYWWCPKAE

WEGSEGLKTLRILYEEVDESEVEVIHVPSPALERKTD SYRYPRTGSKNPKIALKLAEFQ  
 . S G K.LRILYEE DESEVE:IHV.SP.LE.R.:DS:RYP:TG:.NPK:...K.:E..  
 TTPSGG-KILRILYEEVDESEVEIIVTSPMLETRRADSFYRYPKTGTANPKVTFKMSIIM

TDSQKGIVSTQEKELVQPFSSLFKVEYIARAGWTRDGKYAWAMFLDRPQQWLQVLVLLPP  
 .D::G:I:::KEL:QPF. LF. VEYIARAGWT :GKYAW::LDR.Q. LQ:VL:.P  
 IDAEGRIIDVIDKELIQPFELFEGVEYIARAGWTPEGKYAWSILDRSQTRLQIVLISP

ALFIPSTENEEQRLASARAVPRNVQPYVVEVTNVWINVHDI FYPFPQSEGEDELCLFLR  
 .LFIP :.: :R . . .:VP :V.P:::YEE.T::WIN:HDIF: FPQS. E:E: F:  
 ELFIPVEDDVMERQRLIESVPDSVTPLIIEETTDIWINIHDI FHVFPQSH-EEIEFIF

ANECKTGPFCHLYKVTAVLKSQGYDWSEPFSPGEDEFKCPIKEEMALTSGEWEVLARHGSK  
 A:ECKTGF HLYK:T::LK.. Y. S. P...:FKCPIKEE:A:TSGEWEVL.RHGS.  
 ASECKTGFRLHYKITSILKESKYKRSSGGLPAPSDFKCPIKEEIAITSGEWEVLGRHGSN

IWVNEETKLIVYFQGT KDTPLEHHLYVVS YEAAAGEIVRLTTPGFSSHSCSMSQNFDMFVSHY  
 I V:E .:LVYF:GTKD:PLEHHLYVVS Y .GE:.RLT. G:SHSC.:SQ: D.F:S.Y  
 IQVDEVRRLVYFEGTKDSPLEHHLYVVS YVNPGEVTRLTDRGYSHSCCISQHCDFFIISKY

SSVSTPPCVHVYKLSGPD DDPLHKQPRFWASMMEAASCPDYPPEIFHFHTRSDVRLYG  
 S: .P CV.:YKLS.P:DDP. K...FWA:::A. PDY.PPEIF.F:::..LYG  
 SNQKNPHCVSLYKLSPPEDDPTCKTKEFWATILDSAGPLPDYTPPEIFSFESTTGFTLYG

< Prolyl oligopeptidase family (underlined) >

MIYKPHALQPGKKHPTVLFVYGGPQVQLVNN SFKGIKYLRLNTLASLG YAVVVIDGRGSC  
 M:YKPH LQPGKK:PTVLF:YGGPQVQLVNN.FKG:KY.RLNTLASLG Y.VVVID.RGSC  
 MLYKPHDLQPGKKYPTVLFYIYGGPQVQLVNNRFGVKYFRLNTLASLG YVVVIDNRGSC



- 11/13 -

Fig. 29 (continued)

< active site: S residue >

QRGLRFEGALKNQMGQVEIEDQVEGLQFVAEKYGFIDLSRVAIHGWSYGGFLSLMGLIHK  
 .RGL:FEGA.K :MGQ:EI:DQVEGLQ::A.:Y.FIDL.RV.IHGWSYGG:LSLM.L::  
 HRGLKFEGAFKYKMGQIEIDDQVEGLQYLASRYDFIDLDRVGIHGWSYGGYLSLMALMQR

PQVFKVAIAGAPVTVWMAYDTGYTERYMDVPENNQHGYEAGSVALHVEKLPNEPNRLLIL  
 ...F:VAIAGAPVT:W: YDTGYTERYM. P:.N:.GY .GSVA:...EK.P:EPNRL:L  
 SDIFRVAIAGAPVTLWIFYDTGYTERYMGHPDQNEQGYLGSVAMQAEPFPSEPNRLLLL

< active site: D residue >      < active site: H residue >

HGFLDENVHFFETNFLVSQIRAGKPYQLQIYPNERHSIRCPESGEHYEVTLLHFLQEYL      892  
 HGFLDENVHF HT:.L:S L:RAGKPY.LQIYP.ERHSIR.PESGEHYE: LLH:LQE L  
 HGFLDENVHFAHTSILLSFLVRAGKPYDLQIYPQERHSIRVPESGEHYELHLLHYLQENL      872

- 12/13 -

Fig. 30

HMMPFAM - alignment of 181 Protein against pfam|hmm|DPPIV\_N\_term  
Dipeptidyl peptidase IV (DPP IV) N-termi -

This hit is scoring at : -72.9; E= 2.4e-07

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```

Q: 132 LSWKQMLDHFQATP-HHGVYSRE---EELLRERKRLGVFGITSYDFHSESGLFLFQASNS
      :::..... :...: E.L : : : : :Y: ..: L. :. N.
H: 1 vtledifsgtfrpkdydinWIssedgeylyqdqdnli.vynyetgkttvllsdefne

LFHCR-----DGGKNGFMVSPMKPL-----EIKTQ-----CSG-----
      . . . . . D . . . : K . . . : T . . . : G
      feasnksYeveiSpDlkyiLlstnyekrWRHSytasYiyDlntgdvepvapeelgdDNY

---PRMDPKICPADPAFFSFNNNDLWVANIETGEERLTFchQGLSNVlddpkSAGVAT
      . . . . . P . . . :F:::L:V.. :G.. :T .G SN : G:..
      llnkIqyatWSPNkGhklaFVRdNNLYvknpsgpaiqiTt..dGksndI....fNGipD

FVIQEEFDRFTG-YWWCPTASwegseglktlRILYEEVDESEVEVIHVPSPALEE-----
      :V::EE. . . .WW.P... :Y...:SEV.VI..P . :.
      WVYEEElstdyAlWWSPDGd.....fLAYlrfnDseVPvievPfytddskveie

---RKTDsYRYPRTGSKNPKIALKLAEFQTdsqGKIVSTqeKELVQPFSSLFKVEYIAR
      :T...YP:.G:.NP:.L :..... G. VST . :. :L . YI..
      dqYpetmeikYPKAGapNPtVkLfvnlad...gasvst..pveiplpanlasgdyYite

AGWTR--DGKYAWAMFLDRPQQWLQLVLLppalfipstenEEQRLASARAVPRNVQpyvV
      . W.. : :A...L:R.Q. L L. :... :. V :N :
      VtWvtmknerla.VqwlRdQnisvlslc.....Dtasssktwnvvnre...h

YEEVTNVWINVHD-IFYPFPQSEGE-DELCFLRANECKTG--FCHLYKVTAVLKSqgydw
      .EE..W:.... : FP . . : :L : : :G : HL ... KS
      ieasetgWvetfnpslpvfpldgl srlkeyYliisd.rdGgkykHlayyeldgks.....

sepfspgedefkcpikeEMALTSGEWEVLarhgsK--IW-VNEETKL VYFQGTKDTPLEH
      E.:ALT.G.WEV. . I V::T.VYF.T:..E.
      .....epialTkGnWEVt.....ngiilgvdsktltvyFlatekgsgR

HLYVVSYEAGE-----IVRLTT--PGFsHSCSMSQNFDMFVSHYSSVSTPPCVHVYK
      HLY :S:... : . . G: :S.S.S.N : : : YS. .P. :.
      hlysislkggetsktClscqldserDCcgy.ySasFSsnakYyiltysGPgvPtigtllhs

LSGP-----DDDPLH--KQPRFWASMMEAAscppdyVPPEIFHFHTR-----SD
      . . . . . D.: . . . . : :P.: F. . .
      skdekveksdlGkvkdkelrtLEdNeaLkkaLknyq.....lPskefkkklklpdDFadg

VRLYGMiYKPHALQP-GKKHPTVLFVYGGPQVQLVNN 683
      :.L...KP..P.KK:P.:FVYGGP Q V..
      itlnyqmikPanFdpLsKKYPvLffvYGGPgSQqVtk 566

```

- 13/13 -

Fig. 31

HMPFAM - alignment of 181\_Protein against pfam|hmm|Peptidase\_S9  
Prolyl oligopeptidase family

This hit is scoring at : 36.8; E= 4.3e-10

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

Q: 692 RLNTLASLGYAVVVIDGRGSCQRGLRFEGALKNQMGQVEIEDQVEGLQFVaEKYGFIDLS

...L . G ...V:D RG. : G ::. A .... : E..D ..... :K.G:... .

H: 10 vasllnhrGgiyAvvdiRGgGeyGqkwheagtrrlkknefnDfiaAAeyl.sklGYtspk

RVAIHGWSYGGFL 764

R:AI.G S GG.L

riaifGgSnGGlL 81

Prolyl endopeptidase family serine proteins. region from residue 734 to  
764. Source: [blocks database]

## SEQUENCE LISTING

<110> Bayer AG

<120> REGULATION OF HUMAN DIPEPTIDYL PEPTIDASE 8

<130> Lio 298

<150> USSN 60/268,863

<151> 2001-02-16

<160> 28

<170> PatentIn version 3.1

<210> 1

<211> 2676

<212> DNA

<213> Homo sapiens

<400> 1  
atgcggaagg ttaagaaact ggcctggac aaggagaaca ccggaagttg gagaagcttc 60  
tcgctgaatt ccgagggggc tgagaggatg gccaccaccg ggacccaac ggccgaccga 120  
ggcgacgcag ccgccacaga tgaccgggcc gcccgcttcc aggtgcagaa gcactcgtgg 180  
gacgggctcc ggagcatcat ccacggcagc cgcaagtact cgggcctcat tgtcaacaag 240  
gcgccccacg acttccagtt tgtgcagaag acggatgagt ctgggccccca ctcccaccgc 300  
ctctactacc tgggaatgcc atatggcagc cgtgagaact ccctcctcta ctccgagatc 360  
cccaagaaag tgcggaagga ggccctgctg ctgctgtcct ggaagcagat gctggaccac 420  
ttccaggcca caccacacca tgggtgtctac tcccagagagg aggagctact gcgggagcgc 480  
aagcgctgg gcgtcttcgg aatcacctct tatgacttcc acagtgagag cggcctcttc 540

ctcttccagg ccagcaatag cctgttccac tgcagggatg gtggcaagaa tggctttatg 600  
gtgtcccccga tgaagccact ggagatcaag actcagtgtt ctgggccacg catggacccc 660  
aaaatctgcc ctgccgaccc tgccttcttc tccttcatca ataacagcga cctgtgggtg 720  
gccaacatcg agacaggcga ggagcggcgg ctgaccttct gccaccaagg tttatccaat 780  
gtcctgggatg accccaagtc tgcgggtgtg gccaccttcg tcatacagga agagtctgac 840  
cgcttcactg ggtactggtg gtgccccaca gcctcctggg aaggttcaga gggcctcaag 900  
acgtgcgaa tcctgtatga ggaagtcgat gagtccgagg tggaggtcac tcacgtcccc 960  
tctcctgcgc tagaagaaag gaagacggac tcgtatcggc accccaggac aggcagcaag 1020  
aatcccaaga ttgccttgaa actggctgag ttccagactg acagccaggg caagatcgtc 1080  
tcgaccaggg agaaggagct ggtgcagccc ttcagctcgc tgttcccgaa ggtggagtac 1140  
atcgccaggg ccgggtggac ccgggatggc aaatacgccg gggccatggt cctggaccgg 1200  
ccccagcagt ggctccagct cgtcctcttc ccccgccg tgttcacccc gagcacagag 1260  
aatgaggagc agcggctagc ctctgccaga gctgtcccca ggaatgtcca gccgtatgtg 1320  
gtgtacgagg aggtcaccaa cgtctggatc aatgttcatg acatcttcta tcccttcccc 1380  
caatcagagg gagaggacga gctctgcttt ctccgcgcca atgaatgcaa gaccggcttc 1440  
tgccatthgt acaaagtac cgcgthtta aaatcccagg gctacgattg gagtgaagccc 1500  
ttcagccccg gggaagatga atttaagtgc cccattaagg aagagatggc tctgaccagc 1560  
ggtgaatggg aggtthtggc gaggcacggt tccaagatct ggggtcaatga ggagaccaag 1620  
ctggtgtact tccagggcac caaggacacg ccgctggagc accacctcta cgtggtcagc 1680  
tatgaggcgg ccggcgagat cgtacgcctc accacgcccg gcttctccca tagctgctcc 1740  
atgagccaga acttcgacat gttcgtcagc cactacagca gcgtgagcac gccgccctgc 1800  
gtgcacgtct acaagctgag cggccccgac gacgaccccc tgcacaagca gcccgccttc 1860  
tgggctagca tgatggaggc agccagctgc ccccggtatt atgttctctc agagatcttc 1920  
catttccaca cgcgctcgga tgtgcggctc tacggcatga tctacaagcc ccacgccttg 1980  
cagccagggg agaagcacc caccgtcctc tttgtatatg gagggcccca ggtgcagctg 2040  
gtgaataact ccttcaaagg catcaagtac ttgcggtcga acacactggc ctccctgggc 2100  
taecgcgtgg ttgtgattga cggcaggggc tcctgtcagc gagggcttcg gttcgaaggg 2160  
gccctgaaaa accaaatggg ccaggtggag atcgaggacc aggtggaggg cctgcagttc 2220  
gtggccgaga agtatggctt cattgacttg agccgagtcg ccatccatgg ctggtcctac 2280  
gggggcttcc tctcgtcat ggggctaata cacaagcccc aggtgttcaa ggtggccatc 2340

```

gcgggtgcc cggtcaccgt ctggatggcc tacgacacag ggtacactga gcgctacatg 2400
gacgtccctg agaacaacca gcacggctat gaggcgggtt ccgtggctct gcacgtggag 2460
aagctgcccc atgagcccaa ccgcttgctt atcctccacg gcttcctgga cgaaaacgtg 2520
cactttttcc acacaaactt cctcgtctcc caactgatcc gagcagggaa accttaccag 2580
ctccagatct accccaacga gagacacagt attcgtgcc ccgagtcggg cgagcactat 2640
gaagtcacgt tgctgcactt tctacaggaa tacctc 2676

```

&lt;210&gt; 2

&lt;211&gt; 892

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 2

```

Met Arg Lys Val Lys Lys Leu Arg Leu Asp Lys Glu Asn Thr Gly Ser
1          5          10          15

Trp Arg Ser Phe Ser Leu Asn Ser Glu Gly Ala Glu Arg Met Ala Thr
20          25          30

Thr Gly Thr Pro Thr Ala Asp Arg Gly Asp Ala Ala Ala Thr Asp Asp
35          40          45

Pro Ala Ala Arg Phe Gln Val Gln Lys His Ser Trp Asp Gly Leu Arg
50          55          60

Ser Ile Ile His Gly Ser Arg Lys Tyr Ser Gly Leu Ile Val Asn Lys
65          70          75          80

Ala Pro His Asp Phe Gln Phe Val Gln Lys Thr Asp Glu Ser Gly Pro
85          90          95

His Ser His Arg Leu Tyr Tyr Leu Gly Met Pro Tyr Gly Ser Arg Glu
100         105         110

Asn Ser Leu Leu Tyr Ser Glu Ile Pro Lys Lys Val Arg Lys Glu Ala
115         120         125

```

4/27

Leu Leu Leu Leu Ser Trp Lys Gln Met Leu Asp His Phe Gln Ala Thr  
 130 135 140

Pro His His Gly Val Tyr Ser Arg Glu Glu Glu Leu Leu Arg Glu Arg  
 145 150 155 160

Lys Arg Leu Gly Val Phe Gly Ile Thr Ser Tyr Asp Phe His Ser Glu  
 165 170 175

Ser Gly Leu Phe Leu Phe Gln Ala Ser Asn Ser Leu Phe His Cys Arg  
 180 185 190

Asp Gly Gly Lys Asn Gly Phe Met Val Ser Pro Met Lys Pro Leu Glu  
 195 200 205

Ile Lys Thr Gln Cys Ser Gly Pro Arg Met Asp Pro Lys Ile Cys Pro  
 210 215 220

Ala Asp Pro Ala Phe Phe Ser Phe Ile Asn Asn Ser Asp Leu Trp Val  
 225 230 235 240

Ala Asn Ile Glu Thr Gly Glu Glu Arg Arg Leu Thr Phe Cys His Gln  
 245 250 255

Gly Leu Ser Asn Val Leu Asp Asp Pro Lys Ser Ala Gly Val Ala Thr  
 260 265 270

Phe Val Ile Gln Glu Glu Phe Asp Arg Phe Thr Gly Tyr Trp Trp Cys  
 275 280 285

Pro Thr Ala Ser Trp Glu Gly Ser Glu Gly Leu Lys Thr Leu Arg Ile  
 290 295 300

Leu Tyr Glu Glu Val Asp Glu Ser Glu Val Glu Val Ile His Val Pro  
 305 310 315 320

Ser Pro Ala Leu Glu Glu Arg Lys Thr Asp Ser Tyr Arg Tyr Pro Arg  
 325 330 335

Thr Gly Ser Lys Asn Pro Lys Ile Ala Leu Lys Leu Ala Glu Phe Gln  
 340 345 350

Thr Asp Ser Gln Gly Lys Ile Val Ser Thr Gln Glu Lys Glu Leu Val  
 355 360 365

Gln Pro Phe Ser Ser Leu Phe Pro Lys Val Glu Tyr Ile Ala Arg Ala  
 370 375 380

Gly Trp Thr Arg Asp Gly Lys Tyr Ala Trp Ala Met Phe Leu Asp Arg  
 385 390 395 400

Pro Gln Gln Trp Leu Gln Leu Val Leu Leu Pro Pro Ala Leu Phe Ile  
 405 410 415

Pro Ser Thr Glu Asn Glu Glu Gln Arg Leu Ala Ser Ala Arg Ala Val  
 420 425 430

Pro Arg Asn Val Gln Pro Tyr Val Val Tyr Glu Glu Val Thr Asn Val  
 435 440 445

Trp Ile Asn Val His Asp Ile Phe Tyr Pro Phe Pro Gln Ser Glu Gly  
 450 455 460

Glu Asp Glu Leu Cys Phe Leu Arg Ala Asn Glu Cys Lys Thr Gly Phe  
 465 470 475 480

Cys His Leu Tyr Lys Val Thr Ala Val Leu Lys Ser Gln Gly Tyr Asp  
 485 490 495

Trp Ser Glu Pro Phe Ser Pro Gly Glu Asp Glu Phe Lys Cys Pro Ile  
 500 505 510

Lys Glu Glu Met Ala Leu Thr Ser Gly Glu Trp Glu Val Leu Ala Arg  
 515 520 525

His Gly Ser Lys Ile Trp Val Asn Glu Glu Thr Lys Leu Val Tyr Phe  
 530 535 540

Gln Gly Thr Lys Asp Thr Pro Leu Glu His His Leu Tyr Val Val Ser  
 545 550 555 560

Tyr Glu Ala Ala Gly Glu Ile Val Arg Leu Thr Thr Pro Gly Phe Ser  
 565 570 575

His Ser Cys Ser Met Ser Gln Asn Phe Asp Met Phe Val Ser His Tyr  
 580 585 590



Ser Ser Val Ser Thr Pro Pro Cys Val His Val Tyr Lys Leu Ser Gly  
 595 600 605

Pro Asp Asp Asp Pro Leu His Lys Gln Pro Arg Phe Trp Ala Ser Met  
 610 615 620

Met Glu Ala Ala Ser Cys Pro Pro Asp Tyr Val Pro Pro Glu Ile Phe  
 625 630 635 640

His Phe His Thr Arg Ser Asp Val Arg Leu Tyr Gly Met Ile Tyr Lys  
 645 650 655

Pro His Ala Leu Gln Pro Gly Lys Lys His Pro Thr Val Leu Phe Val  
 660 665 670

Tyr Gly Gly Pro Gln Val Gln Leu Val Asn Asn Ser Phe Lys Gly Ile  
 675 680 685

Lys Tyr Leu Arg Leu Asn Thr Leu Ala Ser Leu Gly Tyr Ala Val Val  
 690 695 700

Val Ile Asp Gly Arg Gly Ser Cys Gln Arg Gly Leu Arg Phe Glu Gly  
 705 710 715 720

Ala Leu Lys Asn Gln Met Gly Gln Val Glu Ile Glu Asp Gln Val Glu  
 725 730 735

Gly Leu Gln Phe Val Ala Glu Lys Tyr Gly Phe Ile Asp Leu Ser Arg  
 740 745 750

Val Ala Ile His Gly Trp Ser Tyr Gly Gly Phe Leu Ser Leu Met Gly  
 755 760 765

Leu Ile His Lys Pro Gln Val Phe Lys Val Ala Ile Ala Gly Ala Pro  
 770 775 780

Val Thr Val Trp Met Ala Tyr Asp Thr Gly Tyr Thr Glu Arg Tyr Met  
 785 790 795 800

Asp Val Pro Glu Asn Asn Gln His Gly Tyr Glu Ala Gly Ser Val Ala  
 805 810 815

Leu His Val Glu Lys Leu Pro Asn Glu Pro Asn Arg Leu Leu Ile Leu  
 820 825 830

His Gly Phe Leu Asp Glu Asn Val His Phe Phe His Thr Asn Phe Leu  
835 840 845

Val Ser Gln Leu Ile Arg Ala Gly Lys Pro Tyr Gln Leu Gln Ile Tyr  
850 855 860

Pro Asn Glu Arg His Ser Ile Arg Cys Pro Glu Ser Gly Glu His Tyr  
865 870 875 880

Glu Val Thr Leu Leu His Phe Leu Gln Glu Tyr Leu  
885 890

<210> 3

<211> 882

<212> PRT

<213> Homo sapiens

<400> 3

Met Ala Ala Ala Met Glu Thr Glu Gln Leu Gly Val Glu Ile Phe Glu  
1 5 10 15

Thr Ala Asp Cys Glu Glu Asn Ile Glu Ser Gln Asp Arg Pro Lys Leu  
20 25 30

Glu Pro Phe Tyr Val Glu Arg Tyr Ser Trp Ser Gln Leu Lys Lys Leu  
35 40 45

Leu Ala Asp Thr Arg Lys Tyr His Gly Tyr Met Met Ala Lys Ala Pro  
50 55 60

His Asp Phe Met Phe Val Lys Arg Asn Asp Pro Asp Gly Pro His Ser  
65 70 75 80

Asp Arg Ile Tyr Tyr Leu Ala Met Ser Gly Glu Asn Arg Glu Asn Thr  
85 90 95

Leu Phe Tyr Ser Glu Ile Pro Lys Thr Ile Asn Arg Ala Ala Val Leu  
100 105 110

Met Leu Ser Trp Lys Pro Leu Leu Asp Leu Phe Gln Ala Thr Leu Asp  
 115 120 125

Tyr Gly Met Tyr Ser Arg Glu Glu Leu Leu Arg Glu Arg Lys Arg  
 130 135 140

Ile Gly Thr Val Gly Ile Ala Ser Tyr Asp Tyr His Gln Gly Ser Gly  
 145 150 155 160

Thr Phe Leu Phe Gln Ala Gly Ser Gly Ile Tyr His Val Lys Asp Gly  
 165 170 175

Gly Pro Gln Gly Phe Thr Gln Gln Pro Leu Arg Pro Asn Leu Val Glu  
 180 185 190

Thr Ser Cys Pro Asn Ile Arg Met Asp Pro Lys Leu Cys Pro Ala Asp  
 195 200 205

Pro Asp Trp Ile Ala Phe Ile His Ser Asn Asp Ile Trp Ile Ser Asn  
 210 215 220

Ile Val Thr Arg Glu Glu Arg Arg Leu Thr Tyr Val His Asn Glu Leu  
 225 230 235 240

Ala Asn Met Glu Glu Asp Ala Arg Ser Ala Gly Val Ala Thr Phe Val  
 245 250 255

Leu Gln Glu Glu Phe Asp Arg Tyr Ser Gly Tyr Trp Trp Cys Pro Lys  
 260 265 270

Ala Glu Thr Thr Pro Ser Gly Gly Lys Ile Leu Arg Ile Leu Tyr Glu  
 275 280 285

Glu Asn Asp Glu Ser Glu Val Glu Ile Ile His Val Thr Ser Pro Met  
 290 295 300

Leu Glu Thr Arg Arg Ala Asp Ser Phe Arg Tyr Pro Lys Thr Gly Thr  
 305 310 315 320

Ala Asn Pro Lys Val Thr Phe Lys Met Ser Glu Ile Met Ile Asp Ala  
 325 330 335

Glu Gly Arg Ile Ile Asp Val Ile Asp Lys Glu Leu Ile Gln Pro Phe  
 340 345 350

Glu Ile Leu Phe Glu Gly Val Glu Tyr Ile Ala Arg Ala Gly Trp Thr  
355 360 365

Pro Glu Gly Lys Tyr Ala Trp Ser Ile Leu Leu Asp Arg Ser Gln Thr  
370 375 380

Arg Leu Gln Ile Val Leu Ile Ser Pro Glu Leu Phe Ile Pro Val Glu  
385 390 395 400

Asp Asp Val Met Glu Arg Gln Arg Leu Ile Glu Ser Val Pro Asp Ser  
405 410 415

Val Thr Pro Leu Ile Ile Tyr Glu Glu Thr Thr Asp Ile Trp Ile Asn  
420 425 430

Ile His Asp Ile Phe His Val Phe Pro Gln Ser His Glu Glu Glu Ile  
435 440 445

Glu Phe Ile Phe Ala Ser Glu Cys Lys Thr Gly Phe Arg His Leu Tyr  
450 455 460

Lys Ile Thr Ser Ile Leu Lys Glu Ser Lys Tyr Lys Arg Ser Ser Gly  
465 470 475 480

Gly Leu Pro Ala Pro Ser Asp Phe Lys Cys Pro Ile Lys Glu Glu Ile  
485 490 495

Ala Ile Thr Ser Gly Glu Trp Glu Val Leu Gly Arg His Gly Ser Asn  
500 505 510

Ile Gln Val Asp Glu Val Arg Arg Leu Val Tyr Phe Glu Gly Thr Lys  
515 520 525

Asp Ser Pro Leu Glu His His Leu Tyr Val Val Ser Tyr Val Asn Pro  
530 535 540

Gly Glu Val Thr Arg Leu Thr Asp Arg Gly Tyr Ser His Ser Cys Cys  
545 550 555 560

Ile Ser Gln His Cys Asp Phe Phe Ile Ser Lys Tyr Ser Asn Gln Lys  
565 570 575

Asn Pro His Cys Val Ser Leu Tyr Lys Leu Ser Ser Pro Glu Asp Asp  
 580 585 590

Pro Thr Cys Lys Thr Lys Glu Phe Trp Ala Thr Ile Leu Asp Ser Ala  
 595 600 605

Gly Pro Leu Pro Asp Tyr Thr Pro Pro Glu Ile Phe Ser Phe Glu Ser  
 610 615 620

Thr Thr Gly Phe Thr Leu Tyr Gly Met Leu Tyr Lys Pro His Asp Leu  
 625 630 635 640

Gln Pro Gly Lys Lys Tyr Pro Thr Val Leu Phe Ile Tyr Gly Gly Pro  
 645 650 655

Gln Val Gln Leu Val Asn Asn Arg Phe Lys Gly Val Lys Tyr Phe Arg  
 660 665 670

Leu Asn Thr Leu Ala Ser Leu Gly Tyr Val Val Val Val Ile Asp Asn  
 675 680 685

Arg Gly Ser Cys His Arg Gly Leu Lys Phe Glu Gly Ala Phe Lys Tyr  
 690 695 700

Lys Met Gly Gln Ile Glu Ile Asp Asp Gln Val Glu Gly Leu Gln Tyr  
 705 710 715 720

Leu Ala Ser Arg Tyr Asp Phe Ile Asp Leu Asp Arg Val Gly Ile His  
 725 730 735

Gly Trp Ser Tyr Gly Gly Tyr Leu Ser Leu Met Ala Leu Met Gln Arg  
 740 745 750

Ser Asp Ile Phe Arg Val Ala Ile Ala Gly Ala Pro Val Thr Leu Trp  
 755 760 765

Ile Phe Tyr Asp Thr Gly Tyr Thr Glu Arg Tyr Met Gly His Pro Asp  
 770 775 780

Gln Asn Glu Gln Gly Tyr Tyr Leu Gly Ser Val Ala Met Gln Ala Glu  
 785 790 795 800

Lys Phe Pro Ser Glu Pro Asn Arg Leu Leu Leu Leu His Gly Phe Leu  
 805 810 815

Asp Glu Asn Val His Phe Ala His Thr Ser Ile Leu Leu Ser Phe Leu  
                   820                                  825                                  830

Val Arg Ala Gly Lys Pro Tyr Asp Leu Gln Ile Tyr Pro Gln Glu Arg  
                   835                                  840                                  845

His Ser Ile Arg Val Pro Glu Ser Gly Glu His Tyr Glu Leu His Leu  
                   850                                  855                                  860

Leu His Tyr Leu Gln Glu Asn Leu Gly Ser Arg Ile Ala Ala Leu Lys  
                   865                                  870                                  875                                  880

Val Ile

<210> 4

<211> 611

<212> DNA

<213> Homo sapiens

<400> 4

gagggttcaga gggctcaaga cgctgcgaat cctgtatgag gaagtcgatg agtccgaggt	60
ggagggtcatt cacgtcccct ctcctgcgct agaagaaagg aagacggact cgtatcggta	120
ccccaggaca ggcagcaaga atcccaagat tgccttgaaa ctggctgagt tccagactga	180
cagccagggc aagatcgtct cgaccagga gaaggagctg gtgcagccct tcagctcgct	240
gttcccgaag gtggagtaca tcgccagggc cgggtggacc cgggatggca aatacgctg	300
ggccatgttc ctggaccggc cccagcagtg gctccagctc gtctctctcc ccccggccct	360
gttcatcccg agcacagaga atgaggagca gcggctagcc tctgccagag ctgtccccag	420
gaatgtccag ccgtatgtgg tgtacgagga ggtcaccaac gtctggatca atgttcatga	480
catctttctat cccttcccc aatcagaggg agaggacgag ctctgctttc tccgcgcaa	540
tgaatgcaag accggcttct gccatttgta caaagtcacc gccgttttaa aatccagggc	600
tacgattgga g	611

<210> 5  
<211> 1080  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<222> (1079)..(1080)  
<223> N=a,t,g or c

<400> 5  
ggggagggttt tggcgaggca cggctccaag ggcaccaagg acacgccgct ggagcaccac 60  
ctctacgtgg tcagctatga ggcggccggc gagatcgtag gcctcaccac gcccggttc 120  
tcccatagct gctccatgag ccagaacttc gacatgttcg tcagccacta cagcagcgtg 180  
agcacgccgc cctgcgtgca cgtctacaag ctgagcggcc ccgacgacga cccctgcac 240  
aagcagcccc gcttctgggc tagcatgatg gaggcagcca gctgcccccc ggattatggt 300  
cctccagaga tcttccattt ccacacgcgc tcggatgtgc ggctctacgg catgatctac 360  
aagccccacg ccttgcagcc agggaagaag cccccaccg tcctctttgt atatggaggc 420  
ccccagggtgc agctggtgaa taactccttc aaaggcatca agtacttgcg gctcaacaca 480  
ctggcctccc tgggctacgc cgtggttggtg attgacggca gggctcctgt cagcgagggc 540  
ttcggttcga aggggccctg aaaaaacaaa tgggccaggt ggagatcgag gaccagggtg 600  
gagggcctgc agttccgtga gcgagaagta tggttcctc cgaactgagc cgagttgcca 660  
tccaatggct gggctctacgg ggacttactc tcgctcatgg gggtaatcca caagcccagg 720  
gggtcagggg caatcgcgag tgcccgggca cggctggatg gccaagaaac agcgtacact 780  
gagcggaagg gacggccccg aaaaccagac aggaaaagg gagtccgggg ccgcgtcagg 840  
gagaaagtgc cacgaagcca cagcggggat cacacagagg cgtgcggaca cgggggggtt 900  
ccacaaaaca cgacacaagg agacgggaaa caccaacaac aataccagga acaatgaggc 960  
gcggggaaac gacacgagaa caacacatgg acgcgacgaa caaagcaatg cgacagcaga 1020  
gacgcgggac gaccaggacg ccagttcgca cggtcgccct cggcgcgcg agcgaccagn 1080

&lt;210&gt; 6

&lt;211&gt; 744

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 6

gctcggatgt gcggtctctac ggcattgatct acaagcccca cgccttgtag ccagagaaga	60
agcaccaccac cgtcctcttt gtatatggag gccccagggt gcagctgggtg aataactcct	120
tcaaaggcat caagtacttg cggctcaaca cactggcctc cctgggctac gccgtgggtg	180
tgattgacgg caggggctcc tgtcagcagag ggcttcgggt cgaaggggccc ctgaaaaacc	240
aaatgggcca ggtggagatc gaggaccagg tggagggcct gcagttcgtg gccgagaagt	300
atggcttcat cgacactgag ccgagttgcc atccatggct ggtcctacgg gggttctctc	360
tcgctcatgg ggctaatacca caagccccag gtgttcaagg tggccatcgc gggtgccccg	420
gtcaccgtct ggatggccta cgacacaggg tacactgagc gctacatgga cgtccctgag	480
aacaaccagc acggctatga ggcgggttcc gtggccctgc acgtggagaa gctgccaat	540
gagcccaacg attgcttata ctccacgggt cctggacgaa aacgtgcact ttttccacac	600
aaattcctcg tctcccaact gatccgagca gggaaacctt acagcgtcgg atcaacccaa	660
gagagaacag tttcgtgcc gagtcggcga catatgagtc gttgatgact tctaaggaaa	720
ctctgacctg cacggacgcc atta	744

&lt;210&gt; 7

&lt;211&gt; 506

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 7

aaagcagagc tcgtcctctc cctctgattg ggggaaggga tagaagatgt catgaacatt	60
gatccagacg ttggtgacct cctcgtacac cacatacggc tggacattcc tggggacagc	120
tctggcagag gctagccgct gctcctcatt ctctgtgctc gggatgaaca gggccggggg	180
gaggaggacg agctggagcc actgctgggg ccggtccagg aacatggccc aggcgtattt	240
gccatcccgg gtccaccgg ccctggcgat gtactccacc ttcgggaaca gcgagctgaa	300



gggctgcacc agctccttct cctgggtcga gacgatcttg ccctggctgt cagtctggaa 360  
ctcagccagt ttcaaggcaa tcttgggatt cttgctgcct gtectgggggt accgatacga 420  
gtccgtcttc ctttcttcta gcgcaggaga ggggacgtga atgacctcca cctcggactc 480  
atcgacttcc tcatacagga ttcgca 506

<210> 8

<211> 753

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (688)..(689)

<223> n=a,t,g or c

<400> 8

attacatcta ttttaaagga aagcaaatat aaacgatcca gtggtgggct gcctgctcca 60  
agtgatttca agtgtcctat caaagaggag atagcaatta ccagtgggtga atgggaagtt 120  
cttggccggc atggatctaa tatccaagtt gatgaagtca gaaggctggt atattttgaa 180  
ggcaccaaag actccccctt agagcatcac ctgtacgtag tcagttacgt aaatcctgga 240  
gaggtgacaa ggctgactga ccgtggctac tcacattctt gctgcatcag tcagcactgt 300  
gacttcttta taagtaagta tagtaaccag aagaatccac actgtgtgtc cctttacaag 360  
ctatcaagtc ctgaagatga cccaacttgc aaaacaaagg aattttgggc caccattttg 420  
gattcagcag gtccctcttc tgactatact cctccagaaa ttttctcttt tgaaagtact 480  
actggattta cattgtatgg gatgtcttac aagcctcatg atctacagcc tggaaagaaa 540  
tatectactg tgetgttcat atatggtggc cctcaggtgc agttgggtgaa taatcggttt 600  
aaaggagtca agtatttccg cttgaatacc ctagcctctc taggttatgt ggttgtagtg 660  
atagacaaca ggggatcctg tcaccgangg cttaaatttg aaggcgcctt taaatataaa 720  
atgggtcaaa tagaaattga cgatcaggtg gga 753

&lt;210&gt; 9

&lt;211&gt; 494

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 9

aaagcagagc tcgtcctctc cctctgattg ggggaaggga tagaagatgt catgaacatt	60
gateccagacg ttggtgacct cctcgtacac cacatacggc tggacattcc tggggacagc	120
tctggcagag gctagccgct gctcctcatt ctctgtgctc gggatgaaca gggccggggg	180
gaggaggacg agctggagcc actgctgggg ccggtccagg aacatggccc aggcgtatatt	240
gccatcccgg gtccacccgg ccctggcgat gtactccacc ttcgggaaca gcgagctgaa	300
gggctgcacc agctccttct cctgggtcga gacgatcttg ccctggctgt cagtctggaa	360
ctcagccagt ttcaaggcaa tcttgggata cttgctgcct gtcctggggg accgatacga	420
gtccgtcttc ctttcttcta gcgcaggaga ggggacgtga atgacctcca cctcggactc	480
atcgacttcc tcat	494

&lt;210&gt; 10

&lt;211&gt; 754

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 10

cggctccaag ggcaccaagg acacgccgct ggagcaccac ctctacgtgg tcagctatga	60
ggcggccggc gagatcgtac gcctcaccac gcccggttc toccaatagct gctccatgag	120
ccagaacttc gacatgttcg tcagccacta cagcagcgtg agcacgccgc cctgcgtgca	180
cgtctacaag ctgagcggcc ccgacgacga cccctgcac aagcagcccc gcttctgggc	240
tagcatgatg gaggcagcca gctgcccccc ggattatgtt cctccagaga tcttccattt	300
ccacacgcgc tcggatgtgc ggctctacgg catgatctac aagccccacg ccttgacgcc	360
aggggaagaag caccaccccg tcctcttttg tatatggagg ccccagggtg cagctgggtga	420
ataactcctt caaaggcatc aagtacttgc ggctcaacac actggcctcc ctgtggctac	480
gccgtggttt gtgattgacg gcaggggctc ctgtcagcga gggcttcggt tcgaaggggc	540

ctgaaaaacc aaatgggcca gtgtggagat cgaggacctg gtggagggcc tgcagttcgt 600  
ggccgagaag tatggcttct cgactgagcc gagttgcctc cttggctgtc ctacgggggc 660  
ttctctcgtc catggggcta tccacagccc gtgtccgtgg ctccgggtgc cggcccttgg 720  
ttgctacgac cggtcctggg ctctgatctg gaac 754

<210> 11

<211> 1169

<212> DNA

<213> Homo sapiens

<400> 11

gcggtgtcg tccccgctc ccgccacttc cggggtcgca gtccccgggca tggagccgcg 60  
accgtgaggc gccgctggac cggggacgac ctgccagtc cggccgcccgc cccacgtccc 120  
ggtctgtgtc ccacgcctgc agctggaatg gaggtctctc ggacccttta gaaggcacc 180  
ctgccctcct gaggtcagct gagcggttaa tgcggaaggc taagaaactg cgcctggaca 240  
aggagaacac cggaagtgg agaagcttct cgtgaattc cgagggggct gagaggatgg 300  
ccaccaccgg gaccccaacg gccgaccgag gcgacgcagc cgccacagat gacccggccg 360  
cccgttcca ggtgcagaag cactcgtggg acgggctccg gagcatcatc cacggcagcc 420  
gcaagtactc gggcctcatt gtcaacaagg cgcgccacga cttccagttt gtgcagaaga 480  
cggatgagtc tgggccccac tcccaccgcc tctactacct gggaatgcc a tatggcagcc 540  
gagagaactc cctcctctac tctgagattc ccaagaaggc ccggaaagag gctctgctgc 600  
tccttgctct ggaagcagat gctggcatca tttccaggcc acgccccaca tgggggteta 660  
ctctcgggag gaggagctgc tgaggggagcg gaaacgcctg ggggctcttc ggcatcacct 720  
cctacgactt ccacagcgag agtggcctct tcctcttcca ggccacacag ccttctccac 780  
tggcgcgacg gcggaagaac ggctcatggg tgtccctatg aaccgtggga attcacaccc 840  
aagggacagg gccccgagtg ggagcgaaaa tctgcccggc gaccatgacg tatatgccgt 900  
caaacaacaa caggcaccgc tgggtgtgcc catcctaaca cgggacgaaa cgagcggacc 960  
ctccgcccac gggatcacgc cgagagacc caactcagtg gggggacact gcagagacaa 1020  
ccaacacgaa tcgtacacgt ggcgcacccc cgcgacacag aagcaacaca gcaacgaaga 1080  
agcagacaca gggcgcaagc cgacctagac agagcagacc gcaggacggt acgagcacia 1140  
acatctgaag accgcagcca ccgcccgc 1169

<210> 12

<211> 1002

<212> DNA

<213> Homo sapiens

<400> 12

gagtcaagta tttcccgctt gaatacccta gcctctctag gttatgtggt tgtagtgata	60
gacaacaggg gatectgtca ccgagggctt aaatttgaag gcgcctttaa atataaaatg	120
ggtcaaataag aaattgacga tcaggtggaa ggactccaat atctagcttc tcgatatgat	180
ttcattgact tagatcgtgt gggcatccac ggctggtcct atggaggata cctctccctg	240
atggcattaa tgcagaggtc agatatcttc agggttgcta ttgctggggc cccagtcact	300
ctgtggatct tctatgatac aggatacacg gaacgttata tgggtcaccc tgaccagaat	360
gaacagggct attacttagg atctgtggcc atgcaagcag aaaagttccc ctctgaacca	420
aatcgtttac tgctcttaca tggtttctctg gatgagaatg tccattttgc acataccagt	480
atattactga gttttttagt gagggctgga aagccatatg atttacagat ctatcctcag	540
gagagacaca gcataagagt tcctgaatcg ggagaacatt atgaactgca tcttttgcac	600
tacettcaag aaaaccttgg atcacgtatt gctgctctaa aagtgatata attttgacct	660
gtgtagaact ctctggtata cactgggtatt tgaccaaagtg aggacgtgta atcgaagcga	720
aaaccagaa atggtcacg ccattgtgtt acctgcattg ttagcattta cttctggaaa	780
attaatgttg gtgccatgcg ggcgcttacg ggtgggggaa attaatactt taaccccatg	840
tgctaatcaa tgtatttctc ggaaaccagt cctagatttc taaaaaaaaa accgagattt	900
cttccagagg gcgcaagggg gccacagggc ttaaaaggct ggaagagacc gtcaatgtgc	960
cagtgtgcaa actctttcgt gagagaattg tatgatagg ag	1002

<210> 13

<211> 969

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (400)..(401)

<223> n=a,t,g or c

<400> 13

ggcttctgcc atgtgtacaa agtcaccgcc gttttaaaat cccagggcta cgatgggagt	60
gagcccttca gccccgggga agatgaattt aagtgcacca ttaaggaaga gatggctctg	120
accagcgggtg aatgggaggt tttggcgagg cacggttcca agatctgggt caatgaggag	180
accaagctgg tgtacttcca gggcaccaag gacacgccgc tggagcacca cctctacgtg	240
gtcagctatg agggggccgg cgagatcgta cgcctcaacc agggccgggt tctcccatag	300
ctggctccat gagccagaac ttcgacatgt tcgtcagcca ctacagcagc gtgagcacgc	360
cgccctgcgt gcacgtctaa caagctggag ctggccccga ngaggaacct ctgcacaagc	420
agcccggtt ctgggctagc atgatggagg cagccagggt ccccggggat tatgttcctc	480
cagagatctt ccatttccac acgcgctcgg atgtgcggtc tacggcatga tctagaagcc	540
caagcttggg gccaggacga agaccaacgg cctcttgat ttggggccca ggtgccgtgg	600
ggctaatacct tcaagggtc aggacatgcg ggtacacaat gggctcctgg gtacgccggg	660
aaggagatgc ggagggggcc ggacaggagg ggcagggcaa agggccgaga aaaattgggc	720
cggaaccgg acctggagag gcaaaattgg cacaacagtg gttaaagcga gaaagggaca	780
ccaagcgagg gacaggagga cctacaagtg ggtgacaacc cgcgtagaga gagaagacgg	840
accacggga tgcaaaagaa cacgagaacc aacaaatcaa aagtagcgag cgcacgaaac	900
aatgcaagcg taagtcttc ttcgctgcc tcgctcgcgc cgtctgtcct gctgttggtg	960
ggagaggag	969

<210> 14

<211> 442

<212> DNA

<213> Homo sapiens

<400> 14  
tggcgaattc ggcacgaggg ggaggttttg gcgaggcacg gctccaaggg caccaaggac 60  
acgccgctgg agcaccacct ctacgtggtc agctatgagg cggccggcga gatcggtacgc 120  
ctcaccacgc cgggettctc ccatagctgc tccatgagcc agaacttcga catgttcgtc 180  
agccactaca gcagcgtgag cagccgccc tgcggtgcacg tctacaagct gagcggcccc 240  
gacgacgacc ccctgcacaa gcagccccgc ttctgggcta gcatgatgga ggcagccagc 300  
tgccccccgg attatgttcc tccagagatc ttccatttcc acacgcgctc ggatgtgcgg 360  
ctctacggca tgatctacaa gcccacgcc ttgcagccag ggaagaagca ccccaccgtc 420  
ctctttgtat atggaggccc cc 442

<210> 15

<211> 735

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (708)..(709)

<223> n=a,t,g or c

<400> 15  
caacagggga tccgtgcacc gagggcttaa atttgaagge gcctttaaat ataaaatggg 60  
tcaaataaaa attgacgac aggtggaagg actccaatat ctacgttctc gatatgattt 120  
cattgactta gatcgtgtgg gcatccacgg ctggtcctat ggaggatacc tctccctgat 180  
ggcattaatg cagaggtcag atatcttcag gggttctatt gctggggccc cagtcactct 240  
gtggatcttc tatgatacag gatacacgga acgttatatg ggtcaccctg accagaatga 300  
acagggctat tacttaggat ctgtggccat gcaagcagaa aagttcccct ctgaaccaa 360  
tcgtttactg ctettacatg gtttcctgga tgagaatgtc cattttgcac ataccagtat 420  
attactgagt tttttagtga gggctggaaa gccatatgat ttacagatct atcctcagga 480  
gagacacagc ataagagttc ctgaatcggg agaacattat gaactgcac ttttgcacta 540  
ccttcaagaa aaccttggat cacgtattgc tgctctaaaa gtgatataat tttgacctgt 600

gtagaactct ctggtataca ctggctatctt aaccaaataa ggaggtttta tcaacagaaa 660  
acacagaatt gatcatcaca ttttgatacc tgccatgtaa catctactnc tgaaaataaa 720  
tgtggtgccca tgcag 735

<210> 16

<211> 888

<212> DNA

<213> Homo sapiens

<400> 16

gtctacgggg gcttctcttc gctcatgggg ctaatccaca agccccaggt gttcaagggtg 60  
gccatcgagg gtgccccggg caccgtctgg atggcctacg acacagggtta cactgagcgc 120  
tacatggacg tccctgagaa caaccagcac ggctatgagg cgggttcctg ggtctctgcac 180  
gtggagaagc tgcccaatga gcccaaccgc ttgcttatcc tccacggctt cctggacgaa 240  
aacgtgcact ttttccacac aaacttcctc gtctcccaac tgatccgagc agggaaacct 300  
taccagctcc agatctacc caacgagaga cacagtatcc gctgccccga gtcggggcgag 360  
cactatgaag tcacgttgct gcactttcta caggaatacc tctgagcctg cccaccggga 420  
gccgccacat cacagcacia gtggctgcag cctccgcggg gaaccaggcg ggagggactg 480  
agtggcccg cggccccagt gaggcacttt gtcccgccca gcgctggcca gccccgagga 540  
gccgtgcct tcaccgccc agccttttta tcctttttta aacgtctctg ggttttatgt 600  
ccgctgcttc ttggttgccg agacagagag atggtggtct cgggccagcc ccttctctc 660  
cccgccttct gggaggagga ggtcacacgc tgatgggcac tggagaggca gaagagactc 720  
agaggagggg gctgcttccg ctggggctcc tgtgacctct tcagtcccc tgcccgggca 780  
ggccgggtccc aggcccaagc tgetttgtcg gtccccggg cggttccac ttgtgttggg 840  
ttggtggggg ggatttccat tttaaagagg gggggggggc ctgttccc 888

<210> 17

<211> 454

<212> DNA

<213> Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (435)..(436)

&lt;223&gt; n=a,t,g or c

&lt;400&gt; 17

```
tgccccacag cctcctggga aggttcagag ggcctcaaga cgctgcgaat cctgtatgag      60
gaagtcgatg agtccgaggt ggaggtcatt cacgtcccct ctctgcgct agaagaaagg      120
aagacggact cgcacgcgta cccaggaca ggcagcaaga atcccaagat tgccttgaaa      180
ctggctgagt tccagactga cagccagggc aagatcgtct cgacccaga gaaagagctg      240
gtgcagccct tcagctcgt gttcccgaag gtggagtaca tcgccagggc cgggtggacc      300
cgggatggca aatacgctg ggccatgttc ctggaccggc ccagcagtg gtcacagctc      360
gtccttcttc ccccgccct gttcatcccg agcacagaga atgaggagca gcggctagcc      420
tctggcgagc tgtcncagga atgtccagcc gtat                                     454
```

&lt;210&gt; 18

&lt;211&gt; 452

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 18

```
gtctttaccc gactttgctc ttttggggtc tttgtcttt tggggtcttt gctcatggtg      60
tcccctatga aaccgctgga aatcaagacc cagtgtctcag ggccccggat ggaccccaaa      120
atctgccttg ccgacctgc cttcttcttc ttcataata acagcgacct gtgggtggcc      180
aacatcgaga caggcgagga gcggcggtg accttctgcc accaagggtt atccaatgtc      240
ctggatgacc ccaagtctgc ggggtgtggc accttcgtca tacaggaaga gttcgaccgc      300
ttcactgggt actggtggtg cccacagcc tcctgggaag gttcagaggg cctcaagacg      360
ctgcgaatcc tgtatgagga agtcgatgag tccgaggtgg aggtcattca cgtccactct      420
cctgcgctag aagaaaggaa gacggactcg ta                                     452
```



<210> 19

<211> 555

<212> DNA

<213> *Homo sapiens*

<400> 19

cgtcgtgtaa tccgtacaat gaacaaacag catattgaac cacacgtgac taacgcgatg	60
agtcgacagc attcgtcagg ctctgctcac cttccccggg gctgaagggc tcaactccaat	120
cgtagccctg ggatttttaa acggcgggtga ctttgtacaa atggcagaag ccggtcttgc	180
aattcattgg cgcggagaaa gcagagctcg tcctctccct ctgattgggg gaagggatag	240
aagatgtcat gaacattgat ccagacgttg gtgacctcct cgtacaccac atacggctgg	300
acattcctgg ggacagctct ggcagaggct agccgctgct cctcattctc tgtgctcggg	360
atgaacaggg ccggggggag gaggacgagc tggagccact gctggggccg gtccaggaac	420
atggcccagg cgtatttgcc atccccgggc caccggccc tggcgatgta ctccaccttc	480
gggaacagcg agctgaaggg ctgcaccagc tccttctcct gggtcgagac gatcttgccc	540
tggctgtcag tctgg	555

<210> 20

<211> 343

<212> DNA

<213> *Homo sapiens*

<400> 20

gatcttgccc tggctgtcag tctggaactc agccagtttc aaggcaatct tgggattctt	60
gctgcctgtc ctgggggtacc gatacgagtc cgtcttcctt tcttctagcg caggagaggg	120
gacgtgaatg acctccacct cggactcctc gacttctcct tacaggattc gcagcgtctt	180
gaggccctct gaaccttccc aggaggctgt ggggcaccac cagtaccag tgaagcggtc	240
gaactcttcc tgtatgacga aggtggccac acccgagac ttggggtcct ccaggacatt	300
ggataaacct tgggtggcaga aggtcagccg ccgctcctcg cct	343

<210> 21

<211> 317

<212> DNA

<213> Homo sapiens

<400> 21

```
aacgggtgat gcggttggtt aagaaactgc gcctggacaa ggagaacacc ggaagttgga      60
gaagcttctc gctgaattcc gagggggctg agaggatggc caccaccggg accccaacgg      120
ccgaccgagg cgacgcagcc gccacagatg acccgggcgc ccgcttcag gtgcagaagc      180
actcgtggga cgggctccgg agcatcatcc acggcagccg caagtactcg ggcctcattg      240
tcaacaaggc gccccacgac ttccagtttg tgcagaagac ggatgagtct gggccccact      300
ccaaccgcat cacgggt                                     317
```

<210> 22

<211> 520

<212> DNA

<213> Homo sapiens

<400> 22

```
ttaatgcgga aggttaagaa actgcgcctg gacaaggaga acaccggaag ttggagaagc      60
ttctcgtga attccgaggg ggctgagagg atggccacca ccgggacccc aacggccgac      120
cgaggcgacg cagccgccac agatgaccgg gccgcccgt tccaggtgca gaagcactcg      180
tgggacgggc tccggagcat catccacggc agccgcaagt actcgggcct cattgtcaac      240
atcgagacag gcgaggagcg gcggtgacc ttctgccacc aaggtttata caatgtcctg      300
gatgacccca agtctgcggg tgtggccacc ttcgtcatac aggaagagtt cgaccgcttc      360
actgggtact ggtggtgccc cacagcctcc tgggaagggt cagagggcct caagacgctg      420
cgaatcctgt atgaggaagt cgatgagtc gaggtggagg tcattcacgt accctctcct      480
gcgctagaag aaaggaagac ggactcgtat cggtagccca                                     520
```

&lt;210&gt; 23

&lt;211&gt; 411

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 23

```
aggctgtggg gcaccaccag taccacgtga agcggtcgaa ctcttcctgt atgacgaacg      60
tggccacacc cgcagacttg gggtcacca ggacattgga taaaccttgg tggcagaagg      120
tcagccgccg ctctcgcct gtctcgcatt ttggccaccc acaggtcgct gttattgatg      180
aaggagaaga aggcaggggc ggcagggcag attttggggg ccatccgggg ccctgagcac      240
tgggtcttga tttccagcgg tttcataggg gacaccatga agccgttctt gccgccgtcg      300
cggcagtggg agaggctgtt gctggcctgg aagaggaagt ggccactctc gctgtggaag      360
tcgtaggagg tgatgccgaa gacccccagg cgtttccgct ccctcagcag c              411
```

&lt;210&gt; 24

&lt;211&gt; 600

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 24

```
ttgtggctca gtgccaataa aactttatit atgaacacag gtggtggggc ggatttgacc      60
tgtgggccgt aggtggccaa cccctgggtt tgctaagggc ctggcacagg ggaggtgctc      120
agtaaactct ctgcctgaat aagccaacag ttgggtgctc taagccctgc cagatcatgc      180
agtcactggg gaaggctggc ttcttgccga tccctgttcc ttctcccgca ggggtgtgctg      240
gctgagtggg gggggccgacc aatgaacaaa cagcatattg aaccacacgt gactaacgcg      300
atgagtcgac agcatctgct aggcctctgct caccctcccc ggggctgaag ggctcactcc      360
aatcgtagcc ctgggatttt aaaacggcgg tgactttgta caaatggcag aagccggtct      420
tgcattcatt ggcgcggaga aagcagagct cgtcctctcc ctctgattgg gggaagggat      480
agaagatgtc atgaacattg atccagacgt tgggtgacct ctccgtacac acatacggct      540
cggacattct ggggacagct ctgcgagagg ctagecgctt gctctcattc tctgtgctcg      600
```

<210> 25  
<211> 645  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<222> (327)..(328)  
<223> n=a,t,g or c

<220>  
<221> misc\_feature  
<222> (349)..(350)  
<223> n=a,t,g or c

<400> 25  
agatcgagga ccaggtggag ggcctgcagt tcgtggccga gaagtatggc ttcacgcacc 60  
tgagccgagt tgccatccat ggctggtect acgggggctt cctctcgtc atgggggctaa 120  
tccacaagcc ccaggtgttc aaggtggcca tcgcgggtgc cccggtcacc gtctggatgg 180  
cctacgacac agggtagacc tgagcgctac atggactgtc ctgagaacaa ccagcacggc 240  
tatgaggcgg gttccgtggc cctgccacgt tgtgagaatg gcttgcacca aatgagcccc 300  
caaacccgc ctttgccttt aatccentcc caacgtggcg tttccctnt gtgttaaccg 360  
gaaaaaacgt tgcaacactt atatttttcg ccaagcgaag gccaaaaaca acggttttcc 420  
ccaattacgg tgattccttt ccccccaag aaccttgga aattccccgg taatgccag 480  
gggggaaaaa ccttaccag ctcccagatc taccccaacg agagacacag tattcgtgc 540  
cccgagtctg ggcgagcact atgaagtcac gttgctgcac ttttccctta acaagggggg 600  
aaaatttaaa ccccttccc ttggaaaggc ccccttgcc ccaa 645

&lt;210&gt; 26

&lt;211&gt; 448

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 26

```
gacataaaac ccaagagcgt ttaaagaaaa ggataaaagg cgtcggggcg gtgaaggcag      60
cggctcctcg gggctggcca gcgctgggcg ggacaaagtg cctcactggg gcccgcgggc      120
cactcagtc ctccegcctg gttccccgcg gaggetgcag ccacttgtgc tgtgatgtgg      180
cggctcccg tgggcaggct cagaggtatt cctgtagaaa gtgcagcaac gtgacttcat      240
agtgtcgcc cgactcgggg cagcgaatac tgtgtctctc gttggggtag atctggagct      300
ggtaaggttt ccctgctcgg atcagttggg agacgaggaa gtttgtgtgg aaaaagtgca      360
cgttttcgtc caggaagccg tggaggataa gcaagcgggt gggctcattg ggcagcttct      420
ccacgtgcag ggccacggaa cccgcctc                                     448
```

&lt;210&gt; 27

&lt;211&gt; 442

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 27

```
gacataaaac ccaagagcgt ttaaaaaaag gataaaaggc gtcggggcgg tgaaggcagc      60
ggctcctcgg ggctggccag cgctgggagg gacaaagtgc ctcactgggg ccccgcgggc      120
actcagtc ctccegcctg ttccegcggg aggctgcagc cacttgtgct gtgatgtggc      180
ggctcccggt gggcaggctc agaggtattc ctgtagaaag tgcagcaacg tgacttcata      240
gtgtctcgcc gactcggggc agcgaatact gtgtctctcg ttggggtaga tctggagctg      300
gtaaggtttc cctgctcgga tcagttggga gacgaggaag ttttgtgtga aaaagtgcac      360
gttttcgtcc aggaagccgt ggaggataag caagcgggtg ggctcattgg gcagcttctc      420
cacgtgcagg gccacggaac cc                                     442
```

<210> 28

<211> 330

<212> DNA

<213> Homo sapiens

<400> 28

tctgggtcgtg cgtggcatgg tgtggacttt ggggtggctcg tgggcttctc cacgtgcagg	60
ccacggaacc cgcctcatag cegtgcctgg tgttctcagt ggacgtccat gtagcgtca	120
gtgtaccctg tgtcgtagtg ccaccagac ggtgaccggg gcaccgcga tggccacctt	180
gaacacctgg ggcttgtgga ttagcccat gagcgagagg aagccccgt aggaccagcc	240
atggatggca actcggctca ggtcgatgaa gccatacttc tcggccacga actgcaggcc	300
ctccacctgg tctcgatct ccaccatgag	330

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP 02/01538

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC 7 C12N9/48 C12N15/12 C12N5/10 C12Q1/68 G01N33/50 C12N15/62				
According to International Patent Classification (IPC) or to both national classification and IPC				
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N				
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched				
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) CHEM ABS Data, EMBL, BIOSIS, EPO-Internal, WPI Data, PAJ, SEQUENCE SEARCH				
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>				
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.		
X	ABBOTT, CATHERINE A. ET AL: "Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8" EUROPEAN JOURNAL OF BIOCHEMISTRY (2000), 267(20), 6140-6150, October 2000 (2000-10), XP002204858 cited in the application the whole document -& DATABASE EMBL 'Online! 6 November 2000 (2000-11-06) "Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, complete cds" Database accession no. AF221634 XP002205003 61.6% identity over 2615 nt with SEQ ID No 1 --- -/-	1-7, 10-12, 18-32, 34-51, 54-56, 61,64-67		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.				
<input checked="" type="checkbox"/> Patent family members are listed in annex.				
<b>* Special categories of cited documents:</b> <table border="0"> <tr> <td>           *A* document defining the general state of the art which is not considered to be of particular relevance            *E* earlier document but published on or after the international filing date            *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)            *O* document referring to an oral disclosure, use, exhibition or other means            *P* document published prior to the international filing date but later than the priority date claimed         </td> <td>           *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention            *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone            *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.            *&amp;* document member of the same patent family         </td> </tr> </table>			*A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *&* document member of the same patent family
*A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *&* document member of the same patent family			
Date of the actual completion of the international search		Date of mailing of the international search report		
19 July 2002		31/07/2002		
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016		Authorized officer  Weikl, M		

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/01538

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL 'Online!  9 February 2001 (2001-02-09)  "Mus musculus adult male testis cDNA,  RIKEN full-length enriched library  clone:4932434F09:homolog to Dipeptidyl  peptidase 8, full insert sequence"  Database accession no. AK016546  XP002204860  63.7% identity over 2499 nt with SEQ ID No  1</p> <p>---</p>	<p>1-7,  10-12,  18-32,  34-51,  54-56,  61,64-67</p>
X	<p>DATABASE GENESEQ 'Online!  24 October 2000 (2000-10-24)  "Human peptidase; HPEP-16; coding  sequence"  Database accession no. AAA37672  XP002205004  97.1% identity over 1520 nt with SEQ ID No  1  the whole document</p> <p>---</p>	<p>1-7,  10-12,  30-32,  36-48</p>
X	<p>DATABASE GENESEQ 'Online!  24 October 2000 (2000-10-24)  "Human peptidase; HPEP-16; protein  sequence"  Database accession no. AAY90299  XP002205005  99.4% identity over 507 aa with SEQ ID No  2  the whole document</p> <p>---</p>	<p>1-7,  10-12,  30-32,  36-48</p>
X	<p>DATABASE GENESEQ 'Online!  8 February 2001 (2001-02-08)  "Human ORFX ORF1390 polynucleotide  sequence SEQ ID No 2779"  Database accession no. AAC75835  XP002205006  95.3% identity over 1570 nt with SEQ ID No  1  the whole document</p> <p>---</p>	<p>1-7,  10-12,  30-32,  36-48</p>
X	<p>DATABASE GENESEQ 'Online!  8 February 2001 (2001-02-08)  "Human ORFX ORF1390 polypeptide sequence  SEQ ID No 2780"  Database accession no. AAB41626  XP002205007  98.9% identity over 660 aa with SEQ ID No  2  the whole document</p> <p>---</p>	<p>1-7,  10-12,  30-32,  36-48</p>
	---	
	-/--	



## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/01538

## C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ABBOTT CATHERINE A ET AL: "Two highly conserved glutamic acid residues in the predicted beta propeller domain of dipeptidyl peptidase IV are required for its enzyme activity." FEBS LETTERS, vol. 458, no. 3, pages 278-284, XP004260275 ISSN: 0014-5793 figure 1	1-7, 10-12, 18-32, 34-51, 54-56, 61,64-67
T	--- SEDO ALEKSI ET AL: "Dipeptidyl peptidase IV-like molecules: Homologous proteins or homologous activities?" BIOCHIMICA ET BIOPHYSICA ACTA, vol. 1550, no. 2, December 2001 (2001-12), pages 107-116, XP002204859 ISSN: 0006-3002 table 1	1-7, 10-12, 18-32, 34-51, 54-56, 61,64-67
P,X	--- WO 01 19866 A (ABBOTT CATHERINE ANNE ;GORELL MARK DOUGLAS (AU); UNIV SYDNEY (AU)) 22 March 2001 (2001-03-22)  the whole document -----	1-7, 10-12, 18-32, 34-51, 54-56, 61,64-67

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

## Continuation of Box I.1

Although claim 59 and claims 37, 46, 54, 55 and 56 (insofar as they relate to in vivo methods) are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition (in those cases where the compound/composition could be searched (see Box I.2)).

## Continuation of Box I.2

Claims Nos.: 8, 9, 13, 14, 15, 16, 17, 33, 52, 53, 57, 58, 59, 60, 62, 63 and 68-71

Present claims 8, 9 (partially) 13, 14, 15 (partially) 16 (partially), 17 (partially) 33, 52, 53, 57, 58, 59, 60, 62, 63 and 68-71 relate to reagents defined by reference to a desirable characteristic or property, namely interaction with DPP8 polynucleotides or polypeptides or modulation of their activity.

The claims cover all products having this characteristic or property, whereas the application does not provide support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for any of these reagents. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the reagents by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has not been carried out for these claims (claims 15, 16 and 17 searched insofar as expression vector is concerned and claim 9 searched insofar as the methods of claims 6 or 7 are concerned).

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/EP 02/01538

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
see FURTHER INFORMATION sheet PCT/ISA/210
2. ☒ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:  
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 02/01538

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 0119866	A	22-03-2001	WO 0119866 A1	22-03-2001
			AU 7394600 A	17-04-2001
			EP 1214344 A1	19-06-2002
<hr/>				